

INVERSE CHARACTERIZATION OF SUBSURFACE CONTAMINANT  
MASS FLUX: A THREE-DIMENSIONAL PHYSICAL AND NUMERICAL  
MODELING STUDY



By

MARK ALAN NEWMAN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2001

## ACKNOWLEDGMENTS

I would like to thank my graduate advisor, Dr. Kirk Hatfield, first, for providing the opportunity for me to pursue my graduate degree; second, for his guidance and sense of humor during all of our work together; and third for his ultimate patience.

I would like to thank Drs. P. Suresh C. Rao, Michael Annable, Wendy Graham, and Jennifer Jacobs for their input and suggestions.

In addition, I would like to thank Drs Kirk Hatfield, Suresh Rao, and Michael Annable for making the research process enjoyable, whether it was in the lab or the backwoods of Canada.

I would like to thank Dr. Joel Hayworth of Applied Research Associates and the numerous others of Armstrong Laboratories at Tyndall Air Force Base, all of whom contributed to my practical education in the scientific methods of search and re-search.

I would like to thank my entire family: my mother, grandparents, brother, sister, brother-in-law, nieces, nephew, and parents-in-law for their love and support.

I would especially like to thank my wife, Tricia, for her love and encouragement, which helped to keep me going, and for lovingly feigning interest in my one-sided discussions of flux and numerical convergence.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS .....	ii
LIST OF TABLES .....	v
LIST OF FIGURES .....	vi
ABSTRACT .....	ix
1. INTRODUCTION .....	1
Modeling Theory .....	3
Conceptual Model .....	5
Contaminant Mass Flux and Transport .....	7
Dissertation Organization .....	10
2. MATHEMATICAL MODEL .....	12
Model Development .....	12
Forward and Inverse Problems .....	18
Optimization .....	20
3. NUMERICAL SOLUTION TECHNIQUES: NONLINEAR OPTIMIZATION .....	24
Gradient-Based Techniques .....	24
Minimum Relative Entropy .....	26
Linear inversion .....	30
Algorithm application .....	36
Random Search Techniques .....	39
Simulated Annealing .....	40
Downhill simplex method .....	45
Simulated annealing with modified downhill simplex .....	49
Annealing schedule .....	50
Algorithm application .....	51
Shuffled Complex Evolution .....	51
Competitive complex evolution .....	57
Algorithm application .....	60

4. PHYSICAL MODEL.....	62
Three-Dimensional Aquifer Model.....	62
Physical and Hydrodynamic Characterization of the System.....	70
Transport characterization experiment .....	73
Experiments for Validating the Mathematical Model .....	80
Multiple Source Ionic Tracer Experiment .....	80
Multiple Source DNAPL Dissolution Experiment .....	85
5. NUMERICAL MODELING RESULTS .....	96
Numerically Simulated Test Problem.....	96
Multiple Source Ionic Tracer Experiment .....	100
Multiple Source DNAPL Experiment.....	119
6. SUMMARY AND CONCLUSIONS .....	141
LIST OF REFERENCES .....	144
BIOGRAPHICAL SKETCH .....	149



## LIST OF TABLES

<u>Table</u>	<u>Page</u>
4-1. Physical, hydraulic, and transport properties of the three-dimensional aquifer model.	66
4-2. System parameters for the transport characterization experiment. ....	74
4-3. System parameters for the multi-source ionic tracer experiment. ....	81
4-4. System parameters for the multi-source DNAPL experiment. ....	89
5-1. Simulated flux values at intermediate flux planes for Day 7. ....	112
5-2. Simulated flux values at intermediate flux planes for Day 8. ....	113
5-3. Simulated flux values at the flux plane for the multiple ionic source experiment. ....	117
5-4. Simulated flux at the intermediate flux planes for day 44 of the multiple-source PCE dissolution experiment. ....	125
5-5. Simulated flux at the intermediate flux planes for day 58 of the multiple-source PCE dissolution experiment. ....	126
5-6. Simulated flux at the source planes for day 44 of the multiple-source PCE dissolution experiment. ....	131
5-7. Simulated flux at the source planes for day 58 of the multiple-source PCE dissolution experiment. ....	132

## LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1-1. Conceptual model diagram and simplifying assumptions.....	6
2-1. Mathematical model diagram and definitions. ....	14
2-2. Schematic of continuous plane source. ....	15
3-1. General minimum relative entropy algorithm .....	37
3-2. General simulated annealing algorithm.....	43
3-3. General description of a simplex in three dimensions.....	46
3-4. Possible outcomes for a step in the downhill simplex method.....	48
3-5. A general SCE algorithm (adapted from Duan et al., 1992). ....	53
3-6. A general description of the array <b>D</b> containing complexes $\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^p$ .....	56
3-7. A general CCE algorithm (Adapted from Duan et al., 1992).....	58
4-1. Front view of three-dimensional aquifer model (Air Force Research Laboratory, Tyndall Air Force Base, Florida).....	63
4-2. Rear view (sampling side) of three-dimensional aquifer model (Air Force Research Laboratory, Tyndall Air Force Base, Florida) .....	64
4-3. Schematic of the multi-sampler distribution along a longitudinal transect of the aquifer model. ....	67
4-4. Schematic of the sample port distribution along a lateral transect ( $x = 50$ cm) of the aquifer model. ....	68
4-5. Two-point sampler assembly.....	69
4-6. Steady unconfined flow through the three-dimensional aquifer model, assuming an essentially linear water table and one-dimensional horizontal flow in the $x$ - direction. ....	72

4-7. Conceptual diagram of transport characterization experiment performed to estimate longitudinal and transverse dispersivities. ....	75
4-8. Chloride breakthrough curves ( $x = 50, 120, \text{ and } 170 \text{ cm}$ ) and the corresponding longitudinal dispersivity values determined using a Levenberg-Marquardt nonlinear parameter estimation routine. ....	76
4-9. Transverse chloride concentration distributions and corresponding transverse dispersivity values determined using a Levenberg-Marquardt nonlinear parameter estimation routine. ....	77
4-10. Source locations (transect $x = 20 \text{ cm}$ ) for multiple source ionic tracer experiment....	82
4-11. Steady state three-dimensional ionic tracer plume distribution for Day 7. ....	83
4-12. Steady state three-dimensional ionic tracer plume distribution for Day 8. ....	84
4-13. DNAPL source locations at transect $x = 50 \text{ cm}$ . ....	90
4-14. DNAPL source locations at transect $x = 80 \text{ cm}$ . ....	91
4-15. Steady state three-dimensional PCE concentration distribution for Day 44. ....	92
4-16. Steady state three-dimensional PCE concentration distribution for Day 58. ....	93
4-17. An excavation photograph showing the excavation grid ( $2 \text{ cm} \times 2 \text{ cm}$ grid cells) and a portion of the DNAPL source zone. ....	94
4-18. Three-dimensional representation of DNAPL source zone based upon digital excavation photographs. (The color shown matches the observed dye intensity in the excavation photographs). ....	95
5-1. Numerically simulated flux grid used as test problem for numerical models. ....	98
5-2. Initial Co input values for numerically simulated test problem. ....	99
5-3. Flux plane discretization for multiple source ionic tracer experiment. ....	102
5-4. A) Flux plane grid and tracer concentrations at $x = 1.7 \text{ m}$ used to segregate search space for ionic tracer experiment. B) Initial Co values used for ionic tracer experiment. ....	104
5-5. SA simulated flux distributions for Day 8. ....	106
5-6. SCE simulated ionic tracer flux distributions for Day 8. ....	107
5-7. MRE simulated flux distributions for Day 8. ....	108
5-8. Observed tracer concentration distributions at the intermediate flux planes for day 8. ....	109



5-9. SA-MRE simulated flux distributions for Day 8.....	115
5-10. Estimated probability density function and 95% confidence interval for a single flux cell from the multiple ionic source tracer experiment.....	116
5-11. SA-MRE simulated flux distributions at the source plane for both Day 7 and 8.....	118
5-12. Model grid with shifted coordinates for multiple source DNAPL experiment. ....	121
5-13. Shifted model grid with observed PCE concentrations at $x = 1.70$ m (Day 44) used to establish initial model values.....	122
5-14. Initial $C_0$ values used for multiple source DNAPL experiment.....	123
5-15. Estimated mean value, probability density function, and 95% confidence interval for a single flux cell from the PCE dissolution experiment.....	127
5-16. SA-MRE simulated PCE distributions at intermediate flux planes (Day 44). ....	128
5-17. SA-MRE simulated PCE distributions at intermediate flux planes (Day 58). ....	129
5-18. SA-MRE simulated flux at the PCE source planes (day 44).....	134
5-19. SA-MRE simulated flux at the PCE source planes (day 58).....	135
5-20. DNAPL source zone distribution recorded during excavation.....	136
5-21. Simulated mean flux values and 95% confidence intervals. All flux values greater than 0.01 are shown sorted in order of descending flux intensity. ....	140

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

INVERSE CHARACTERIZATION OF SUBSURFACE CONTAMINANT  
MASS FLUX: A THREE-DIMENSIONAL PHYSICAL AND NUMERICAL  
MODELING STUDY

By

Mark Alan Newman

December 2001

Chairman: Kirk Hatfield

Cochairman: P. Suresh C. Rao

Major Department: Civil and Coastal Engineering

Most field investigations leave subsurface sources poorly defined. Consequently, contaminant distributions and fluxes estimated using forward modeling techniques contain uncertainties due to poorly characterized source morphology and ill-defined hydrogeochemical properties of the aquifer. The primary goal of this study was to develop a method for using observed down-gradient contaminant concentrations in order to estimate the magnitude and distribution of contaminant mass flux through an arbitrary plane located at or near the source zone. A secondary goal was that the model provide some statement of the uncertainty associated with the estimated contaminant mass flux values.

In order to estimate the magnitude and spatial distribution of contaminant mass flux through a plane, the problem was considered in an optimization framework. Three numerical optimization techniques were applied, two random search techniques



(simulated annealing and shuffled complex evolution) and one gradient-based technique (minimum relative entropy). To test the capabilities of the flux plane model and the numerical solution techniques, ionic tracer and NAPL dissolution experiments were performed in a three-dimensional aquifer model designed to simulate flow through an unconfined sand aquifer. Results demonstrate that the random search techniques provide similar estimates for flux magnitude and spatial distribution, with simulated annealing being more computationally efficient. Neither of the random search techniques is capable of providing an estimate of the uncertainty associated with the simulated flux values. In contrast, the method of minimum relative entropy, because it is a gradient-based technique, is not initially as effective at estimating flux magnitude and spatial distribution. But, once in the neighborhood of the optimal solution, it is an excellent tool for inferring mass flux probability density functions, expected values, and confidence limits.

A coupled simulated annealing and minimum relative entropy solution technique was developed in order to take advantage of the robust solution capabilities of simulated annealing and the uncertainty estimation capabilities of minimum relative entropy. The coupled technique performed better than each of the independent methods, although only slightly better than simulated annealing. But, the flux probability density functions and confidence intervals provided by the coupled technique would not have been available from an independent simulated annealing algorithm and they are more accurate than if provided by an independent minimum relative entropy algorithm.

## CHAPTER 1 INTRODUCTION

Groundwater contamination is a worldwide problem with the list of possible contaminants seeming endless. As with any problem of this magnitude the best method for remediation is constantly a topic of debate. Pump and treat, vapor extraction, cosolvent flushing, and natural attenuation—just to name a few—all have their champions and detractors. However, before the debate over remediation strategies can begin for a specified site, the procedure of primary importance is to accurately assess the magnitude and extent of contamination. Then based upon this assessment future impacts must be considered.

The validity of any contamination risk assessment depends upon adequate characterization of the contaminant source and the aquifer hydrogeological properties. Unfortunately, most field investigations leave subsurface sources poorly defined. This is often due to a lack of resources, but more importantly there are physical limitations on the standard intrusive methods for characterizing contaminant sources such as soil coring and cone penetrometer methods. The development of technologies such as partitioning tracers and interfacial tracers offer in situ characterization alternatives to the traditional intrusive techniques (Jin et al., 1995; Annable et al., 1998a, b). However, application of in situ tracer techniques still require some general knowledge of the contaminant source location and distribution.

The purpose of this study was to develop a tool capable of assisting with



contamination risk assessments by estimating the magnitude and spatial distribution of pollutants within contaminated groundwater systems. Due to the complex processes involved in the dissolution of water-soluble contaminants within a porous media, such as residual zones of nonaqueous phase liquids (NAPLs), determining the true location and distribution of a contaminant source represents a daunting task. Being able to make a statement about the uncertainty associated with the estimated contaminant mass and distribution is even more difficult. With the eventual goal being the ability to provide an accurate estimate for contaminant mass and distribution, a first step in solving such a problem may be to consider the magnitude and spatial distribution of contaminant mass flux through an arbitrary flux plane. By estimating the distribution of contaminant mass flux through a plane, information about the contaminant mass distribution up gradient of the plane can be surmised. This information can then be used to guide further characterization studies, remediation efforts, and impact assessments.

The tool envisioned was a model, more specifically a numerical model, capable of satisfying three goals. The primary goal was to develop a method for using observed down-gradient contaminant concentrations in order to estimate the magnitude and distribution of contaminant mass flux through an arbitrary plane at or near the source zone. The secondary goal was that the model provide some statement of the uncertainty associated with the estimated contaminant mass flux values. The tertiary goal was for the model to be developed as a means for assisting with initial site assessments; this means that the model should be easily applied while requiring minimal input data (observed contaminant concentrations, pore water velocity, and dispersivity values). As more

information becomes available, it can be incorporated into the model, but a large amount of requisite data should not be necessary for initial application of the model.

With the primary objectives of the study established the next step was to develop a model capable of meeting these goals. In order to discuss the development of the model it is first necessary to briefly discuss modeling theory.

### Modeling Theory

A model can be defined as a simplified version of a real system that approximates the response of the real system to applied stresses. Typically, a real system is far too complex to be reproduced precisely, and a simplified version is necessary in order to predict how the system may respond to specified conditions and stresses. The simplified version (model) is established as a set of simplifying assumptions that express our understanding of the nature of the system and its behavior (Bear and Verruijt 1987).

When developing a model, the first step is to establish a conceptual model of the system of interest. Usually, the conceptual model consists of a diagram and simplifying assumptions that reduce the real system to a simplified version and facilitate solution of the objective problem. Once the conceptual model is complete it is used to design subsequent more complex models.

In this study both physical and mathematical models were utilized. Physical models are simplified versions of the real system that are functional and provide the opportunity to perform physical experiments in order to simulate the responses of the system to applied conditions and stresses. There is typically a scale difference between the physical model and the real system; this must be considered in the simplifying assumptions and when analyzing the experimental results.



Mathematical models represent the real system as a system of simplifying assumptions and mathematical equations developed from the conceptual model. Solving the model equations provides estimates for the system responses. Mathematical models can be further classified based upon their method of solution: analytical or numerical. Analytical solutions are preferable when possible, because they represent a direct solution to the system of equations and are usually very easy to apply. However, in many cases analytical solutions do not exist, and if they do it is sometimes at the cost of oversimplifying the system. Numerical solutions are developed by discretizing the system into small elements. Typically, the partial differential equations involved in the mathematical description of the system are replaced by discrete equations representing the system response within each model element. The elements are then considered as independent systems that share information at their boundaries and the response of the entire system is the combined response of the smaller elements. Often times a mathematical model is referred to by its method of solution, for instance a mathematical model that is solved numerically may be referred to simply as a numerical model. This is mentioned simply to distinguish between the mathematical models developed in Chapter 2 and the resulting “numerical models” that are produced by applying the numerical solution techniques presented in Chapter 3.

Both analytical and numerical models are usually implemented using computers. In the past 10 to 20 years, the trend has been to move toward numerical solutions due to the lack of analytical solutions for many systems of interest, or the over-simplification involved when an analytical solution does exist. However, numerical solutions are far more computationally involved, and because the simulated system must be discretized



according to the system boundaries, numerical models are usually problem specific.

Although analytical solutions may be considered “over-simplified” in many cases, the ease of application and lower computational expense can make them an excellent tool for initial investigation efforts.

### Conceptual Model

Given the previous discussion of modeling theory, the first step in developing a model is to establish a conceptual model. As stated previously, the primary objective of this study is to develop a model capable of using observed down-gradient contaminant concentrations in order to estimate the magnitude and distribution of contaminant mass flux through an arbitrary plane at or near a source zone. The conceptual model developed to consider this primary model objective is presented in Figure 1-1. The system of interest is an unconfined homogeneous sand aquifer in which all transport is taking place within the saturated zone. It is assumed that a zone of water-soluble contaminant has been introduced to the aquifer and that the contaminant is being transported in the direction of the natural groundwater gradient. The groundwater flow is assumed uniform and steady. It is also assumed that contaminant concentrations can be measured at multiple locations down gradient of the contaminant source zone. The flux plane is established so that it is perpendicular to the direction of groundwater flow. The conceptual model presented in Figure 1-1 was used to develop the mathematical and physical models presented in Chapters 2 and 4 respectively.

With the conceptual model established, the next step is to develop a mathematical representation of the system. In order to do this we will need to define mass flux and develop the advective-dispersive contaminant transport equation.

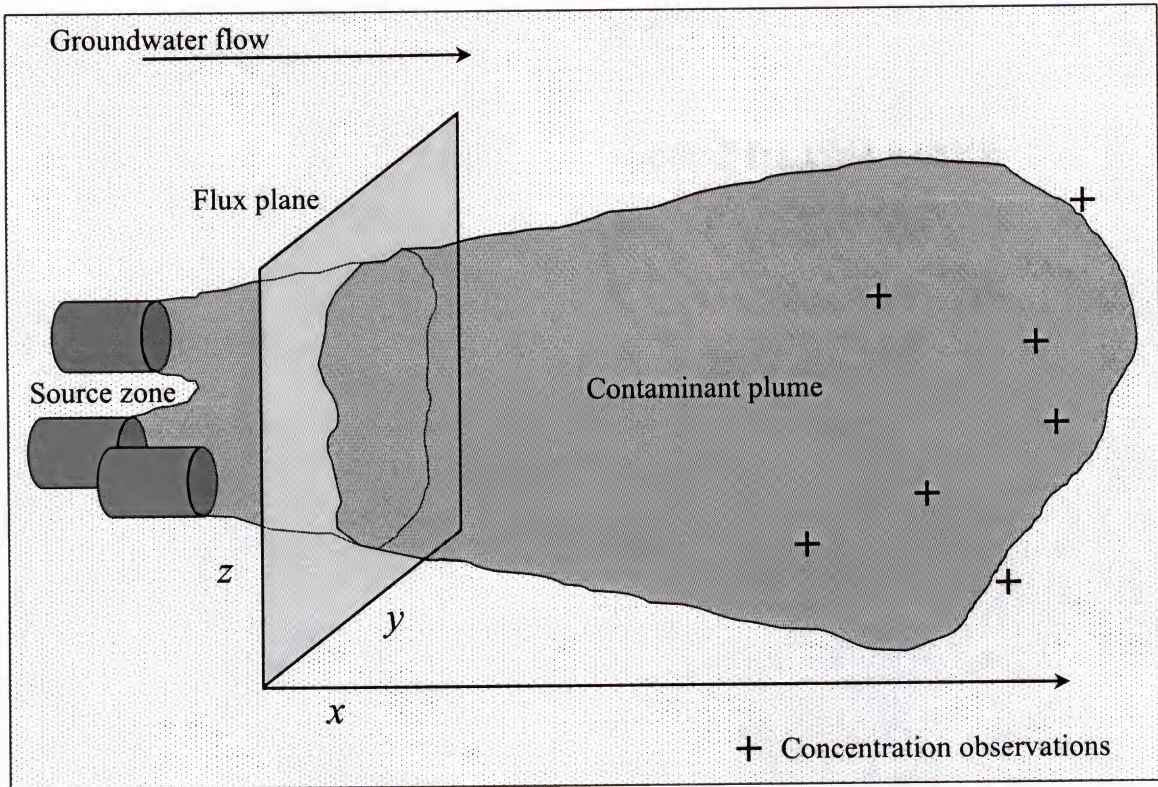


Figure 1-1. Conceptual model diagram and simplifying assumptions.

Conceptual model assumptions:

1. The system is an unconfined sand aquifer
2. The principle axes of the porous medium coincide with the coordinate axes  
 $(D_{xy} = D_{yx} = D_{xz} = D_{zx} = D_{yz} = D_{zy} = 0)$
3. The porous medium is homogeneous and anisotropic  $(D_{xx} \neq D_{yy} \neq D_{zz})$
4. Flow only occurs along the x-axis  $(q_y = q_z = 0)$
5. Groundwater flow is assumed to be uniform  $\left(\frac{\partial q_x}{\partial x} = 0\right)$  and steady  $\left(\frac{\partial q_x}{\partial t} = 0\right)$
6. All transport occurs in the saturated zone  $(\theta = n)$
7. The flux plane is located perpendicular to direction of groundwater flow



### Contaminant Mass Flux and Transport

From the study of physics, flux is defined as the rate of flow of fluid, particles, or energy through a given surface (Wolfson and Pasachoff 1990). In terms of groundwater flow and contaminant transport, flux is used to quantify the mass of water and or contaminants flowing through a specified cross-sectional area during a given period of time. There are several standard notations used when discussing mass flux. In order to maintain consistent notation throughout all mathematical developments in this study, contaminant mass flux will be denoted as  $\dot{m}$ . Based upon the general definition above, the units associated with mass flux can be determined as

$$\dot{m} = \frac{\text{mass}}{\text{unit area} \cdot \text{time}} = \left[ \frac{\text{M}}{\text{L}^2 \cdot \text{T}} \right] \quad (1-1)$$

where the terms M, L, and T represent the base units of mass, length, and time respectively.

The flux of water-soluble contaminants within a water-saturated porous medium has two components: advective flux and dispersive flux

$$\dot{m} = \dot{m}_A + \dot{m}_D \quad (1-2)$$

where  $\dot{m}_A$  is the advective component, and  $\dot{m}_D$  is the dispersive component.

The advective flux represents the transport of water-soluble contaminants induced by groundwater flow, independent of diffusion or mixing. In three-dimensions the general advective flux is

$$\dot{m}_A = \mathbf{q}C \quad (1-3)$$

where  $C$  is the contaminant concentration with units  $[\text{M}/\text{L}^3]$  and  $\mathbf{q}$  is the specific discharge vector (or Darcy velocity) defined as

$$\mathbf{q} = (q_x \hat{\mathbf{i}} + q_y \hat{\mathbf{j}} + q_z \hat{\mathbf{k}})^T \quad (1-4)$$

Each component of the specific discharge  $q_x$ ,  $q_y$ ,  $q_z$  has the appropriate units of velocity [L/T]. Using general matrix notation adapted from Liggett's *Fluid Mechanics* (Liggett 1994), the terms  $\hat{\mathbf{i}}$ ,  $\hat{\mathbf{j}}$ ,  $\hat{\mathbf{k}}$  are the unit vectors along the coordinate axes  $x$ ,  $y$ ,  $z$  respectively. The superscript index **T** represents the vector transpose and indicates that for the purpose of matrix operations  $\mathbf{q}$  is actually a column vector.

The dispersive flux represents the spread of contaminants due to molecular diffusion and mechanical dispersion, and is described based upon an analog to Fick's first law of diffusion. In three-dimensions the general dispersive flux is

$$\mathbf{m}_D = -\theta \mathbf{D} \cdot \nabla C = -\theta \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \cdot \left( \frac{\partial C}{\partial x} \hat{\mathbf{i}} + \frac{\partial C}{\partial y} \hat{\mathbf{j}} + \frac{\partial C}{\partial z} \hat{\mathbf{k}} \right)^T \quad (1-5)$$

where  $\theta$  is the soil moisture content,  $\mathbf{D}$  is a tensor containing the hydrodynamic dispersion coefficients with respect to the coordinate axes  $x$ ,  $y$ ,  $z$ , and the term  $\nabla C$  represents the concentration gradient, or rate of change of  $C$  with respect to each of the coordinate axes:  $\frac{\partial C}{\partial x}$ ,  $\frac{\partial C}{\partial y}$ ,  $\frac{\partial C}{\partial z}$ . The negative sign indicates that the dispersive flux  $\mathbf{m}_D$  is

in the opposite direction of the concentration gradient  $\nabla C$ . This means that dispersive flux (or spreading) occurs in the direction of decreasing concentration.

The general equation for advective-dispersive mass flux can now be represented as



$$\dot{\mathbf{m}} = C \left( q_x \hat{\mathbf{i}} + q_y \hat{\mathbf{j}} + q_z \hat{\mathbf{k}} \right)^T - \theta \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \bullet \left( \frac{\partial C}{\partial x} \hat{\mathbf{i}} + \frac{\partial C}{\partial y} \hat{\mathbf{j}} + \frac{\partial C}{\partial z} \hat{\mathbf{k}} \right)^T \quad (1-6)$$

From equation (1-6) it can be seen that in the general case, mass flux represents a vector quantity meaning that it has components along each of the primary axes. Equation (1-6) can be simplified by applying assumption 2 from the conceptual model (Figure 1-1: Assumption 2. The principle axes of the porous media coincide with the coordinate axes), and the advective-dispersive mass flux reduces to

$$\dot{\mathbf{m}} = C \left( q_x \hat{\mathbf{i}} + q_y \hat{\mathbf{j}} + q_z \hat{\mathbf{k}} \right)^T - \theta \left( D_{xx} \frac{\partial C}{\partial x} \hat{\mathbf{i}} + D_{yy} \frac{\partial C}{\partial y} \hat{\mathbf{j}} + D_{zz} \frac{\partial C}{\partial z} \hat{\mathbf{k}} \right)^T \quad (1-7)$$

Given the above expression for contaminant mass flux, the equation describing contaminant transport (the rate of change of contaminant concentration within a system) is derived by considering conservation of mass (continuity). The general relationship of mass continuity states that for a given system (Domenico and Schwartz 1990)

$$\text{temporal changes in mass storage} = \text{net spatial changes in mass flux}$$

or mathematically

$$\frac{\partial(\theta C)}{\partial t} = -\nabla \cdot \dot{\mathbf{m}} = -\left( \frac{\partial \dot{m}}{\partial x} + \frac{\partial \dot{m}}{\partial y} + \frac{\partial \dot{m}}{\partial z} \right) \quad (1-8)$$

Note that in equation (1-8) the matrix operation  $\nabla \cdot \dot{\mathbf{m}}$  represents the divergence of  $\dot{\mathbf{m}}$ , which produces a scalar quantity, and matrix notation is no longer necessary.

By considering assumptions 3-5 of the conceptual model (Figure 1-1), mass continuity provides the following transport equation:



$$\frac{\partial(\theta C)}{\partial t} = -q_x \frac{\partial C}{\partial x} + \theta \left( D_{xx} \frac{\partial^2 C}{\partial x^2} + D_{yy} \frac{\partial^2 C}{\partial y^2} + D_{zz} \frac{\partial^2 C}{\partial z^2} \right) \quad (1-9)$$

Dividing both sides of equation (1-9) by  $\theta$ , and applying assumption 6 from the conceptual model ( $\theta = n$ , where  $n$  is the porosity) further reduces the transport equation to

$$\frac{\partial C}{\partial t} = -v_x \frac{\partial C}{\partial x} + D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} + D_z \frac{\partial^2 C}{\partial z^2} \quad (1-10)$$

where  $v_x$  is the pore water velocity defined as  $\frac{q_x}{n}$  with the units [L/T]. Due to simplifying assumptions 2 and 3, the redundant subscript was dropped from each of the dispersion coefficients. Equation (1-10) represents the final form of the advective-dispersive contaminant transport equation developed based upon the conceptual model (Figure 1-1). This equation will be used to develop the mathematical model presented in Chapter 2.

### Dissertation Organization

Based upon the conceptual problem (Figure 1-1) and contaminant transport equation (1-10) developed in this chapter, a mathematical model is presented in Chapter 2. The mathematical model presents a method for relating observed contaminant concentrations to elemental flux cells within an arbitrary flux plane. This relationship is dependent upon a transfer function that is an analytical solution to the contaminant transport equation (1-10). Once the mathematical model and transfer function are established, solution of the resulting system of equations is considered as an optimization problem.

Chapter 3 presents three nonlinear optimization techniques that were independently applied to solve the mathematical model. One gradient based technique (minimum relative entropy) and two random search techniques (simulated annealing and shuffled complex evolution) were used.

Chapter 4 presents the physical model that was constructed in order to test the capabilities of the mathematical model and solution techniques. Laboratory investigations were performed in a large, three-dimensional physical model configured to simulate flow through a surficial sand aquifer. The laboratory experiments consisted of an initial non-reactive tracer experiment performed to estimate dispersivity values, followed by a multiple-point-source non-reactive tracer experiment, and a DNAPL (dense nonaqueous phase liquid) multiple-source dissolution experiment.

Chapter 5 provides the results of the numerical modeling study and compares the abilities of the solution techniques to estimate the magnitude and distribution of contaminant mass flux values produced in the physical modeling experiments. Also discussed is a hybrid method of solution utilizing both simulated annealing and the method of minimum relative entropy capable of inversely simulating (optimizing) the magnitude and distribution of contaminant mass flux values while providing an estimate of the uncertainty corresponding to each of the simulated values. Chapter 6 presents the final conclusions and a discussion of the possible applications for the hybrid model solution.



## CHAPTER 2 MATHEMATICAL MODEL

### Model Development

As discussed in Chapter 1, the process of developing a mathematical model begins with a conceptual model (Figure 1-1). The mathematical model is created by representing the conceptual model in terms of mathematical equations that will facilitate solution of the problem established by the goals of the investigation. For this study the modeling goals were as follows:

1. Develop a method for using observed down-gradient contaminant concentrations in order to estimate the magnitude and distribution of contaminant mass flux through an arbitrary plane at or near the source zone.
2. Provide some statement of the uncertainty associated with the estimated contaminant mass flux values.
3. Develop the model as a tool for initial site assessments; it should be easily applied while requiring minimal input data.

Considering the first goal, it is necessary to establish a mathematical relationship between the contaminant mass flux through the plane and the  $j$  downgradient concentration observations (Figure 1-1). In order to do this the flux plane is divided into  $N$  rectangular elements each having a flux component  $m_n$  (Figure 2-1). The total mass flux through the plane is then the sum of the  $N$  elemental fluxes, and the concentration at location  $j$ , ( $C_j$ ) has a component contributed from each of the elemental fluxes  $m_n$ .

$$C_j = \sum_{n=1}^N g_{jn} m_n \quad (2-1)$$

where  $g_{jn}$  is a transfer function relating the flux through element  $n$ ,  $m_n$  to the concentration at location  $j$ ,  $C_j$ . Equation (2-1) is used to produce simulated down gradient concentration values for comparison with observed values and is the basis of the mathematical model (Figure 2-1).

After establishing equation (2-1), the next step is to determine the form of the transfer function  $g_{jn}$ . With the flux plane divided into  $N$  flux elements, it is assumed that each flux element can be simulated as a continuous plane source. Based upon the work of Hunt (1978), Domenico and Robbins (1985) developed an analytical solution for three-dimensional transport from a continuous plane source by solving the advective-dispersive transport equation presented in Chapter 1 (Equation 1-10). The Domenico and Robbins solution is shown below as modified for this investigation.

(2-2)

$$C = \frac{C_o}{8} \operatorname{erfc} \left[ \frac{x - vt}{2(\alpha_x vt)} \right] \left\{ \operatorname{erf} \left[ \frac{y + b}{2(\alpha_y x)^{1/2}} \right] - \operatorname{erf} \left[ \frac{y - b}{2(\alpha_y x)^{1/2}} \right] \right\} \left\{ \operatorname{erf} \left[ \frac{z + d}{2(\alpha_z x)^{1/2}} \right] - \operatorname{erf} \left[ \frac{z - d}{2(\alpha_z x)^{1/2}} \right] \right\}$$

Assuming that the origin for the coordinate axes is located at the center of the plane source, the coordinates  $x$ ,  $y$ ,  $z$  represent the location of the downgradient concentration  $C$  (Figure 2-2). The dimensions  $b$  and  $d$  represent the half-width and half-height of the plane source.  $C_o$  is the source concentration,  $v$  is the pore water velocity,  $t$  represents the elapsed time, and  $\alpha_x$ ,  $\alpha_y$ ,  $\alpha_z$  represent the dispersivity components along the  $x$ ,  $y$ , and  $z$ -axes.



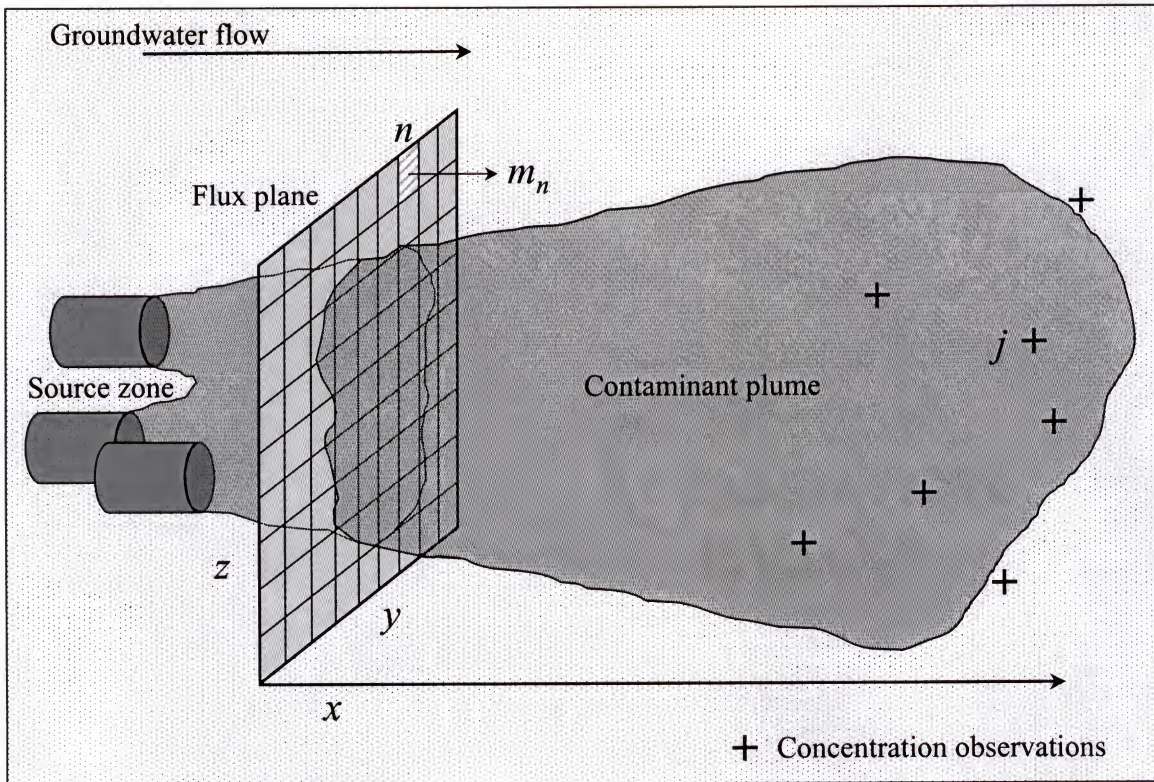


Figure 2-1. Mathematical model diagram and definitions.

Mathematical model definitions:

1. The flux plane is divided into  $N$  elemental flux cells, where the flux through each cell is  $m_n$ , for  $n = (1, 2, \dots, N)$ .
2. There are  $M$  concentration observations located down gradient of the flux plane. The observations are denoted using the index  $j = (1, 2, \dots, M)$ .
3. The simulated concentration  $C_j$  at location  $j$  can be represented as a function of the flux  $m_n$  through each of the flux plane elements:

$$C_j = \sum_{n=1}^N g_{jn} m_n$$

where  $g_{jn}$  is a transfer function relating the flux through element  $n$ ,  $m_n$  to the concentration at location  $j$ ,  $C_j$ .



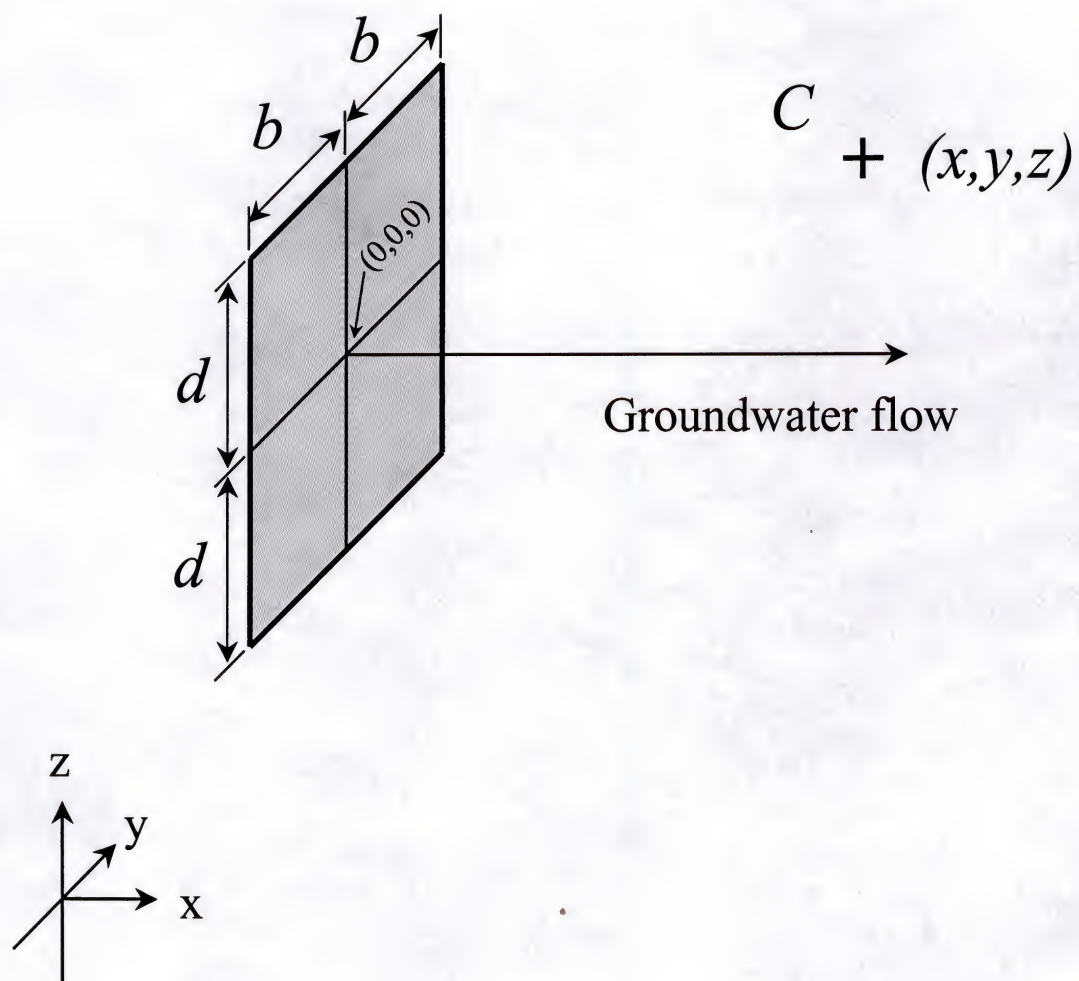


Figure 2-2. Schematic of continuous plane source.

Dispersivity is a parameter used to quantify dispersion within a porous medium.

The dispersivities are related to the dispersion coefficients by the following relationships:

$$D_x = \alpha_x v \quad D_y = \alpha_y v \quad D_z = \alpha_z v \quad (2-3)$$

where  $D_x, D_y, D_z$  are the dispersion coefficients and  $v$  is the pore water velocity. The units for  $\alpha_x, \alpha_y, \alpha_z$  are length [L] and they are characteristic properties of a medium (Bear 1972). The original form of equation (2-2) presented by Domenico and Robbins (1985) was developed in terms of dispersion coefficients  $D_x, D_y, D_z$ . For this study the equation was rewritten in terms of dispersivity in order to reduce the number of terms.

Given the analytical solution for three-dimensional transport from a continuous plane source (2-2) and assuming steady-state conditions:

$$\text{As } t \rightarrow \infty; \quad \frac{x - vt}{2(\alpha_x vt)^{1/2}} \rightarrow -\infty; \quad \text{erfc}(-\infty) \rightarrow 2 \quad (2-4)$$

a transfer function  $g_{jn}$  relating the concentration at location  $j$  to flux intensity through element  $n$  can be formulated as

$$g_{jn} = \frac{1}{4q_x} \left\{ \text{erf} \left[ \frac{y'_n + b}{2(\alpha_y x'_n)^{1/2}} \right] - \text{erf} \left[ \frac{y'_n - b}{2(\alpha_y x'_n)^{1/2}} \right] \right\} \left\{ \text{erf} \left[ \frac{z'_n + d}{2(\alpha_z x'_n)^{1/2}} \right] - \text{erf} \left[ \frac{z'_n - d}{2(\alpha_z x'_n)^{1/2}} \right] \right\} \quad (2-5)$$

where  $x'_n, y'_n, z'_n$  are the relative coordinates defined as

$$\begin{aligned} x'_n &= x_{\text{location},j} - x_{\text{element},n} \\ y'_n &= y_{\text{location},j} - y_{\text{element},n} \\ z'_n &= z_{\text{location},j} - z_{\text{element},n} \end{aligned} \quad (2-6)$$

Inspection of equation (2-5) shows that the steady-state assumption removes the longitudinal dispersivity  $\alpha_x$  from the transfer function. This implies that once steady-state



conditions are achieved the lateral and vertical extents of the plume are determined primarily by the transverse dispersivities  $\alpha_y$  and  $\alpha_z$ .

The specific discharge (Darcy velocity)  $q_x$  appears in equation (2-5) in order to allow the transfer function to relate concentration  $C_j$  to flux intensity  $m_n$ . This is done by assuming that at a given location  $j$ , the mass flux can be estimated as the advective flux component  $q_x C_j$  (1-3). By incorporating  $q_x$  into  $g_{jn}$  the transfer function is assigned the units of [T/L] and equation (2-1) provides a direct relationship between downgradient concentrations and flux intensities at the flux plane. If mass flux values were measured directly, the  $q_x$  term can be removed and  $g_{jn}$  would relate downgradient flux values to flux intensity at the plane.

With the transfer function defined (2-5), equation (2-1) is complete and the problem to be solved can be stated as

- |                  |  |
|------------------|--|
| <b>Given</b>     | Measured values for the characteristic system parameters:<br>Specific discharge, $q_x$<br>Dispersivities, $\alpha_x, \alpha_y, \alpha_z$<br>Flux element dimensions, $b$ and $d$             |
| <b>Determine</b> | What values of flux intensity $m_n$ will produce simulated downgradient concentration values $C_j = \sum_{n=1}^N g_{jn} m_n$ that best match the observed concentration values $C_{obs,j}$ . |

Essentially, the goal of the problem is to determine the best  $m_n$  values based upon observed concentrations, which suggests the problem should be solved within an optimization framework. Optimization represents the solution of an inverse problem. Numerous methods of solution exist for inverse problems, each having its own strengths and weaknesses. Selection of appropriate solution techniques requires accurate

characterization of the equations to be solved and a general understanding of the relationship between forward and inverse problems.

### Forward and Inverse Problems

The topic of inverse groundwater problems has been discussed in depth in the literature (Ahlfeld et al., 1986; Wagner and Gorelick 1987; Freyberg 1988; Carrera et al., 1989; Wagner and Gorelick 1989; Sun 1994; McLaughlin and Townley 1996). McLaughlin (1996) and Sun (1994) provide good discussions of the general groundwater problem, while the others discuss the inverse problem more specifically in terms of parameter estimation and optimization.

The processes of optimization and parameter estimation are commonly described as solving an inverse problem. In order to numerically simulate a physical system (such as groundwater flow through a sand aquifer), two problems must be solved, the forward problem and the inverse problem. The purpose of solving the forward problem is to predict the responses of a system to applied stresses or changes in system conditions. The purpose of solving the inverse problem is to estimate unknown system parameters so that observed system responses can be matched or closely reproduced. Typically, the inverse problem must be solved first in order to establish appropriate model structure and parameters (Sun 1994). Then the forward problem is solved in order to predict system responses.

As stated above, solution of the inverse problem is oftentimes a means to an end—system parameters are estimated by fitting observed conditions, and then the calibrated parameters are applied in a forward model for predictive purposes. In some cases, solution of the inverse problem provides the desired information. However, it should be noted that all inverse problems are based upon the corresponding forward



problem formulation, and solution of the inverse problem usually requires successive iterations of the forward problem.

The terms parameter estimation and optimization are sometimes used interchangeably. But typically, the term parameter estimation is used when discussing the solution of an inverse problem that is solved in order to provide parameter estimates that can then be used with a forward model to predict system responses. An example would be the process of estimating aquifer transmissivity values which best fit observed aquifer hydraulic head values, and then applying the estimated transmissivity values with a forward model to predict head values in response to future pumping rates. The term optimization is usually used to discuss the solution of an inverse problem from which the optimal parameters or system conditions are the desired result. An example would be the process of optimizing the individual pumping rates of wells within a municipal well field in order to yield a desired production rate while imposing minimal reduction of hydraulic head within the aquifer.

The problem statement for this study describes a case in which the solution of the inverse problem produces the desired results. Equation (2-1) represents a forward problem—a method for simulating concentration values  $C_j$  based upon flux intensities  $m_n$ . However, the desired results are the  $m_n$  values that will minimize the difference between the simulated values  $C_j$  and observed concentration values  $C_{obs,j}$ . In general terms, the goal is to estimate system parameters or conditions based upon observed system responses—thus an inverse problem. The inverse problem is solved through an iterative process in which the  $m_n$  values are altered and after each change in  $m_n$  values, a forward simulation is performed using equation (2-1) to determine the corresponding simulated concentration values  $C_j$ . Then the simulated values are compared to the

observed concentrations  $C_{obs,j}$ . This process is continued until it is determined that the difference between the simulated and observed values can be reduced no more. The final  $m_n$  values are then considered the optimal flux values.

### Optimization

The goal of an optimization problem is to find the combination of system parameters (independent variables) that optimize a system response or quantity, possibly subject to restrictions on the allowed parameter ranges. The quantity to be optimized is the objective function; the parameters that may be changed while searching for the optimal solution are called the control or decision variables; and the restrictions on allowed parameter values are the constraints.

A general optimization problem can be stated as

$$\begin{aligned} &\text{minimize } f(\mathbf{x}), \quad \mathbf{x} = (x_1, x_2, x_3, \dots, x_n)^T \\ &\text{subject to } c_i(\mathbf{x}) = 0, \quad i = 1, 2, 3, \dots, m \\ &\quad \quad \quad c_j(\mathbf{x}) \geq 0, \quad j = 1, 2, 3, \dots, p \end{aligned} \tag{2-7}$$

where  $f(\mathbf{x})$  is the objective function,  $\mathbf{x}$  is the column vector of  $n$  independent variables (control variables),  $c_i(\mathbf{x}) = 0$  represent equality constraints, and  $c_j(\mathbf{x}) \geq 0$  represent inequality constraints.

In order to cast the problem statement for this study in an optimization framework we must first define an objective function. Because the goal is to minimize the difference between simulated and observed concentration values, the objective function must quantify these differences. By representing the difference between observed concentration values and simulated concentration values as the sum of the squared differences the objective function can be expressed as



$$f = \sum_{j=1}^M \left( C_{obs,j} - \sum_{n=1}^N g_{jn} m_n \right)^2 \quad (2-8)$$

where  $M$  is the number of observations,  $N$  is the number of flux elements,  $C_{obs,j}$  is the observed concentration value at location  $j$ , and  $\sum_{n=1}^N g_{jn} m_n$  represents the simulated concentration value at location  $j$  from equation (2-1). The problem statement can now be represented as

$$\begin{aligned} \text{minimize } f &= \sum_{j=1}^M \left( C_{obs,j} - \sum_{n=1}^N g_{jn} m_n \right)^2 \\ \text{subject to } m_n &\geq L \\ m_n &\leq U \end{aligned} \quad (2-9)$$

where  $L$  represents the lower limit of the flux values, assumed to be 0, and  $U$  is the upper limit used to bracket the feasible solution space and reduce unnecessary calculations. The upper limit  $U$  is assigned based upon prior knowledge of the contaminant of interest. For instance, if the water solubility limit is known, an upper limit for contaminant mass flux can be estimated as the product of the specific discharge and the solubility limit. If no prior information is available, care must be taken to assign an upper limit that is high enough to include all possible flux levels.

There are numerous numerical techniques for the purpose of solving optimization problems such as the one presented in (2-9). Typically, most algorithms are only appropriate for certain types of problems. For this reason it is important to characterize a problem in order to identify the best solution technique(s). Optimization problems are classified according to the mathematical characteristics of the objective function, constraints, and control variables.

Usually, the most important characteristic is the nature of the objective function: linear or nonlinear. If a general function  $f(x_1, x_2, \dots, x_n)$  is not linear with respect to each of its independent variables, it is said to be nonlinear. Inspection of equation (2-5) indicates that the transfer function  $g_{jn}$  is nonlinear, and it follows that the objective function  $f$  in (2-9) is also nonlinear.

The objective function is also characterized based upon the problem formulation—are there restrictions or constraints on the objective function? If the problem is subject to constraints then it is said to be constrained, if not the problem is said to be unconstrained. Typically, unconstrained problems are more readily solved, and for that reason, one of the more common methods for solving constrained optimization problems is to recast the problem so that the objective function is unconstrained. It will be seen that with gradient-based techniques, such as minimum relative entropy, this is a necessary step in model development. However, with random search techniques, such as simulated annealing and shuffled complex evolution, system constraints are easily incorporated into the algorithms. It can be seen from (2-9) that the problem of interest is constrained by the feasibility requirements (upper and lower limits) imposed on the flux values  $m_n$ .

Control variables are characterized in two ways: based upon the number (univariate or multivariate) and type (discrete or continuous) of variables. For the multivariate case the problem can consist of numerous variables of different type. The possible combinations of variables for a multivariate problem are continuous, discrete (integer), mixed integer (both continuous and discrete variables), and combinatorial. In the case of this study there are multiple control variables ( $m_n$ ) each of which is continuous real number.



Based upon the criteria for classifying optimization problems, the problem presented in equation (2-9) requires solution of a nonlinear constrained continuous multivariate optimization problem. With the mathematical model established (2-1) and the problem statement characterized in terms of an optimization framework (2-9), methods of solution can now be considered. The key criteria being that a nonlinear method capable of dealing with multiple variables is required. Three techniques were selected for investigation: one gradient-based technique (minimum relative entropy) and two random search techniques (simulated annealing and shuffled complex evolution). Minimum relative entropy was considered because of its unique ability to provide estimates for the pdf of the simulated parameters, while the random search techniques were considered for their robust search capabilities. Each of the methods are discussed in detail in Chapter 3.

### CHAPTER 3

#### NUMERICAL SOLUTION TECHNIQUES: NONLINEAR OPTIMIZATION

As discussed in Chapter 2 the intent of this study was to consider contaminant mass flux estimation as a nonlinear optimization problem. There is a large body of literature dedicated to the topic of nonlinear optimization and there are numerous methods to choose from. For the purpose of this discussion nonlinear optimization techniques were categorized into two main groups: gradient-based techniques and random search techniques. For this study one gradient-based technique and two random search techniques were investigated. The gradient-based technique investigated was minimum relative entropy (MRE) and the random search techniques were simulated annealing (SA) and shuffled complex evolution (SCE). The intent was to evaluate the independent capabilities of each method and then consider a hybrid method.

#### Gradient-Based Techniques

As the name implies, gradient-based optimization techniques are dependent upon calculating the gradient (first-derivative of a multivariate function) and in some cases, the higher order derivatives of the objective function. Typically, the method utilizes the gradient to determine the direction to search for an optimal value. When a minimum is desired, the method stipulates that all moves must be downhill, or in the direction of decreasing gradient. In some methods the higher order derivatives are utilized to help determine the magnitude and direction of each optimization step. With almost all gradient-based techniques the gradient and higher order derivatives are estimated using



the Taylor series approximation. For an  $N$ -dimensional system with an initial point  $\mathbf{x} = (x_1, x_2, x_3, \dots, x_N)$ , the Taylor series estimate of an arbitrary function  $f(\mathbf{x})$  evaluated at some location  $(\mathbf{x} + \Delta\mathbf{x})$  is given below

$$f(\mathbf{x} + \Delta\mathbf{x}) \approx f(\mathbf{x}) + \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \Delta\mathbf{x} + \frac{1}{2} \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^2} (\Delta\mathbf{x})^2 + \frac{1}{3} \frac{\partial^3 f(\mathbf{x})}{\partial \mathbf{x}^3} (\Delta\mathbf{x})^3 + \dots \quad (3-1)$$

where  $\Delta\mathbf{x}$  is also an  $N$ -dimensional vector  $(\Delta x_1, \Delta x_2, \Delta x_3, \dots, \Delta x_N)$ . Assuming all higher order derivatives to be negligible and considering only the gradient, the discrete form of (3-1) is

$$f(\mathbf{x} + \Delta\mathbf{x}) \approx f(\mathbf{x}) + \frac{\partial f(x_i)}{\partial x_i} \Delta x_i \quad (3-2)$$

This discrete form of the Taylor series is typically used to determine the search direction between each step of a gradient-based optimization algorithm.

The key points to acknowledge when considering gradient-based techniques are that they require the calculation of objective function derivatives, and that they are only capable of moving downhill (in the direction of decreasing gradient). By incorporating gradient calculations, the resulting algorithms are usually very efficient at finding a local minimum when the gradient is steep (because the gradient allows directional searches toward the minimum). However, in cases where the minimum exists in an area with a shallow gradient the method can become inefficient as convergence is approached. Also, by limiting the search to only downhill movements, some gradient-based techniques have a tendency to get trapped in local minima. Another drawback is that if the objective function is extremely nonlinear it is often difficult if not impossible to perform the gradient calculations.

### Minimum Relative Entropy

Minimum relative entropy (MRE) is a gradient-based optimization technique capable of using observation data to infer probability density functions and expected values for unknown model parameters. The term minimal relative entropy is used because the procedure is based upon minimizing the entropy between an unknown pdf and some prior estimate of the pdf. The application of minimum relative entropy to solve a general linear inverse problem was presented by Woodbury and Ulrych (1996) and was based upon the work of Shore and Johnson (1981). The following development was adapted from two papers (Woodbury and Ulrych 1993; Woodbury and Ulrych 1996). Some corrections were necessary and additional elements were provided in hopes of clarifying the procedure.

Provided a mathematical model of a system, the general concept of MRE is to inversely determine an estimate for the pdf of each unknown independent variable based upon an apparent prior pdf inferred from observed dependent variables. The mathematical basis of the procedure is the general relationship for determining the expected value of an arbitrary function of a random variable  $\mathbf{x}$  given that the pdf of  $\mathbf{x}$  is known. In terms of a multidimensional system, the concept is developed below.

For an  $N$ -dimensional system of interest, let  $\mathbf{x}$  be a possible state where  $\mathbf{x}$  is an  $N$ -dimensional vector:

$$\mathbf{x} = (x_1, x_2, x_3, \dots, x_N)^T \quad (3-3)$$

Each element of  $\mathbf{x}$  represents an unknown independent random variable. Using the notation of Woodbury and Ulrych (1993), assume that  $q^+(\mathbf{x})$  is the multivariate



probability density function (pdf) of  $\mathbf{x}$ . The standard definition of a pdf requires that the pdf must be positive for all values of  $\mathbf{x}$ :

$$q^+(\mathbf{x}) \geq 0 \quad (3-4)$$

and that the sum of all probable values of  $\mathbf{x}$  must be equal to one:

$$\int q^+(\mathbf{x}) d\mathbf{x} = 1 \quad (3-5)$$

The integration in equation (3-5) represents a multiple integral (one integral for each of the  $N$  variables in  $\mathbf{x}$ ).

Now that the pdf of the independent random variable  $\mathbf{x}$  has been defined, assume that a dependent random variable exists in the form of a general function  $f(\mathbf{x})$ . The expected (mean) value of the dependent random variable  $f(\mathbf{x})$  is

$$E[f(\mathbf{x})] = \int f(\mathbf{x}) q^+(\mathbf{x}) d\mathbf{x} = \bar{f} \quad (3-6)$$

where  $\bar{f}$  is a scalar quantity representing the mean value of  $f(\mathbf{x})$ .

Equation (3-6) was introduced above in the framework of a forward problem—given a known pdf  $q^+(\mathbf{x})$ , determine the expected value of the dependent random variable  $f(\mathbf{x})$ . Now let us consider the corresponding inverse problem where  $q^+(\mathbf{x})$  represents the pdf of unknown model parameters. Suppose that  $f(\mathbf{x})$  is a measurable quantity, and that a value for  $\bar{f}$  can be estimated based upon prior information in the form of observed values. Knowing  $f(\mathbf{x})$  and  $\bar{f}$ , equation (3-6) can be used to infer a prior estimate for the unknown pdf  $q^+(\mathbf{x})$ . For the case of multiple observations, equation 3-6 takes the form (Woodbury and Ulrych 1993).

$$E[f_j(\mathbf{x})] = \int_{-\infty}^{\infty} f_j(\mathbf{x}) q^+(\mathbf{x}) d\mathbf{x} = \bar{f}_j ; \quad j = 1, 2, \dots, M \quad (3-7)$$

where  $\bar{f}_j$  represents the observed (mean) value at location  $j$  and  $M$  is the total number of observations.

Because we are dealing with continuous random variables, we cannot calculate the exact or “true” pdf  $q^+(\mathbf{x})$ . The best we can do is attempt to establish an estimate  $q(\mathbf{x})$  for the true pdf  $q^+(\mathbf{x})$ . Within an optimization framework the premise is to minimize the difference between a model simulated estimate  $q(\mathbf{x})$  and the prior estimate  $p(\mathbf{x})$  inferred from observation data (equation 3-7). Where  $q(\mathbf{x})$  is referred to as the posterior estimate of the true pdf  $q^+(\mathbf{x})$ .

As with all optimization problems, in order to minimize the difference between observed and simulated values, an efficient method for quantifying the difference has to be established. There are numerous mathematical expressions used to represent the difference between two quantities of interest, one of the simplest being the sum of the squared differences. The choice of measure for representing the difference is typically based upon the type of data and the desired result. Relative entropy is a unique measure for comparing observed and simulated values, because it considers the relative uncertainty.

Entropy  $H$ , is a term from the study of thermodynamics that is used to indicate the disorder or randomness of a system. When applied to optimization, it is a measure of the uncertainty involved with the random variables. By minimizing the relative entropy between two pdfs it is assumed that the uncertainty between the pdfs is also reduced (Kapur, 1992). The standard relative entropy equation in terms of  $q$  and  $p$  is (Woodbury and Ulrych 1993)



$$H(q, p) = \int q(\mathbf{x}) \ln \left[ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right] d\mathbf{x} \quad (3-8)$$

where  $H(q, p)$  is read as, the entropy of  $q$  relative to  $p$ .

Now that a measure for quantifying the difference between prior and posterior estimates has been established, the optimization problem is to minimize the relative entropy of the posterior estimate  $q(\mathbf{x})$  relative to the prior estimate  $p(\mathbf{x})$  (3-8) subject to the unity constraint from the definition of a pdf (3-5) and the expected value constraint based upon observation data (3-7):

$$\text{Minimize } \int q(\mathbf{x}) \ln \left[ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right] d\mathbf{x} \quad (3-9)$$

$$\text{such that } \int q(\mathbf{x}) d\mathbf{x} = 1 \quad (3-10)$$

$$\text{and } \int f_j(\mathbf{x}) q(\mathbf{x}) d\mathbf{x} = \bar{f}_j; \quad j = 1, 2, \dots, M \quad (3-11)$$

Note, that because there are a finite number of observations, a discrete sum over all the observed values ( $j = 1, 2, \dots, M$ ) is implied in equation (3-11).

In order to solve the optimization problem established in equations (3-9), (3-10), and (3-11) the method of Lagrangean optimization is applied to transform the constrained problem to an unconstrained problem. This is done by rewriting the constraint equations so that they are equal to zero, multiplying each constraint equation by an arbitrary Lagrangean multiplier, and adding these terms to the initial constrained objective function (3-9). The resulting unconstrained objective function is (Corrected from Woodbury and Ulrych (1993)).

$$\Phi = \int q(\mathbf{x}) \ln \left[ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right] d\mathbf{x} + \mu \left[ \int q(\mathbf{x}) d\mathbf{x} - 1 \right] + \sum_{j=1}^M \lambda_j \left[ \int f_j(\mathbf{x}) q(\mathbf{x}) d\mathbf{x} - \bar{f}_j \right] \quad (3-12)$$

where  $\mu$  and  $\lambda_j$  are Lagrangean multipliers. The problem is now to minimize  $\Phi$  with respect to  $q(\mathbf{x})$  (evaluate  $\frac{\partial \Phi}{\partial q(\mathbf{x})} = 0$ ). Solving the partial derivative results in the following equation,

$$\frac{\partial \Phi}{\partial q(\mathbf{x})} = \int \left[ \ln \left[ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right] + 1 + \mu + \sum_{j=1}^M \lambda_j f_j(\mathbf{x}) \right] d\mathbf{x} = 0 \quad (3-13)$$

In order for the above integral to be equal to zero, the combination of all terms within the integral must be zero, therefore

$$\ln \left[ \frac{q(\mathbf{x})}{p(\mathbf{x})} \right] + 1 + \mu + \sum_{j=1}^M \lambda_j f_j(\mathbf{x}) = 0 \quad (3-14)$$

which can be solved for the pdf of interest  $q(\mathbf{x})$  (Woodbury and Ulrych 1993).

$$q(\mathbf{x}) = p(\mathbf{x}) \exp \left[ -1 - \mu - \sum_{j=1}^M \lambda_j f_j(\mathbf{x}) \right] \quad (3-15)$$

Equation 3-15 represents the posterior estimate  $q(\mathbf{x})$  in terms of the prior estimate  $p(\mathbf{x})$ . This estimate of the pdf is one of the unique features of the method of minimum relative entropy. However, the pdf alone does not provide an indication of the optimal values for the independent model parameters. To complete the inverse problem the posterior estimate of the pdf  $q(\mathbf{m})$  must be used to calculate expected values for the model parameters. A method for developing the expected values that is directly applicable to the conceptual problem introduced in Chapter 2 is that of linear inversion as presented by Woodbury and Ulrych (1996).

#### Linear inversion

Assume a discrete linear system exists such that

$$\mathbf{d} = \mathbf{g} \mathbf{m} \quad (3-16)$$



where  $\mathbf{d}$  is an array containing known observation data,  $\mathbf{m}$  is a vector of model simulated values for the unknown control parameters, and  $\mathbf{g}$  is an array of known transfer functions that relate  $\mathbf{d}$  to  $\mathbf{m}$ . Although the unknown control variables are continuous random variables and the transfer functions may be nonlinear, the system represented in equation 3-16 is discrete and linear. This is because the number of observations and the number of control parameters is finite and the relationship between  $\mathbf{d}$  and  $\mathbf{m}$  is linear with  $\mathbf{g}$  acting as a constant of proportionality.

The general value for a single observation at point  $j$  is (Woodbury and Ulrych 1996)

$$d_j = \sum_{n=1}^N g_{jn} m_n \quad (3-17)$$

where  $j$  is the index for the number of observations  $M$  ( $j = 1, 2, \dots, M$ ), and  $n$  is the index for the number of unknown control parameters  $N$  ( $n = 1, 2, \dots, N$ ).

If we apply the same notation as the previous section, we can define  $q(\mathbf{m})$  as the posterior estimate for the pdf of  $\mathbf{m}$ , then by applying equation (3-6), the relationship for determining the pdf of a general function of a random variable that has a known pdf, it can be shown that the expected value of  $d_j$  is

$$E[d_j] = \int_M d_j q(\mathbf{m}) d\mathbf{m} = \bar{d}_j \quad (3-18)$$

which upon substitution of (3-17) leads to the following result (Woodbury and Ulrych 1996).

$$\int_M q(\mathbf{m}) \left[ \sum_{n=1}^N g_{jn} m_n \right] d\mathbf{m} = \bar{d}_j \quad (3-19)$$

Equation 3-19 represents the expected or mean observed value  $d_j$  in terms of the unknown model parameters  $\mathbf{m}$ , the posterior estimate of the pdf  $q(\mathbf{m})$ , and the transfer functions  $\mathbf{g}$ .

Because all optimization problems must start with an initial guess for each of the unknown parameters, we must establish a set (or vector) of  $N$  initial values  $\mathbf{S}$ , where each  $S_n$  is the corresponding initial value for  $m_n$ . Then, a method must be developed for relating each of the model simulated parameters  $m_n$ , to the initial values  $S_n$ , and the posterior estimate of the pdf  $q(\mathbf{m})$ . Because the initial values represent prior information it can be assumed that the  $S_n$  values correspond to the expected values of the prior pdf estimate  $p(\mathbf{m})$ .

$$\int_M m_n p(\mathbf{m}) d\mathbf{m} = S_n ; \quad n = 1, 2, \dots, N \quad (3-20)$$

But, what is the form of the pdf  $p(\mathbf{m})$ ? With most inverse problems, as with the case of the conceptual model presented in Chapter 2, the only statement we can make about the unknown model parameters is that they must fall within some range of reasonable values. This range of reasonable values is usually determined based upon previous knowledge of the parameters and system configuration. Lacking any other information, the range of reasonable values is simply represented as a lower bound  $L$ , and upper bound  $U$ . With the lower and upper bounds as the only information available, the best pdf for comparison is a uniform distribution over the interval  $(L, U)$  (Woodbury and Urych 1993). The uniform distribution in terms of the model simulated parameters  $\mathbf{m}$  is shown below.



$$b(\mathbf{m}) = \left\{ \begin{array}{ll} \frac{1}{U_n - L_n} & L_n \leq m_n \leq U_n \\ 0 & \text{otherwise} \end{array} \right\} \quad (3-21)$$

By applying the method of minimum relative entropy, we can establish the form of the prior estimate of the pdf by minimizing the entropy of  $p(\mathbf{m})$  relative to the uniform distribution  $b(\mathbf{m})$ .

$$\text{Minimize } \int p(\mathbf{m}) \ln \left[ \frac{p(\mathbf{m})}{b(\mathbf{m})} \right] d\mathbf{m} \quad (3-22)$$

$$\text{such that } \int p(\mathbf{m}) d\mathbf{m} = 1 \quad (3-23)$$

$$\text{and } \int_M m_n p(\mathbf{m}) d\mathbf{m} = S_n; \quad n = 1, 2, \dots, N \quad (3-24)$$

Utilizing Lagrangean optimization as before and assuming that the lower limit is zero ( $L = 0$ ) the prior estimate of the pdf is shown below. The final Woodbury and Ulrych (1993) solution is correct, but subtle corrections in the development were necessary.

$$p(\mathbf{m}) = \prod_{n=1}^N \frac{-\beta_n}{\exp(\beta_n U) - 1} \exp(-\beta_n m_n) \quad \text{for } L = 0 \quad (3-25)$$

The operator  $\Pi$  represents the iterated product, defined as  $\prod_{n=1}^N a_n = (a_1 \cdot a_2 \cdot \dots \cdot a_N)$ . The

$\beta_n$  terms are Lagrangean multipliers. By substituting equation 3-25 into equation 3-24 and integrating, the resulting expression for  $S_n$  is

$$S_n = \frac{\exp(-\beta_n U) \beta_n U + \exp(-\beta_n U) - 1}{\beta_n (\exp(-\beta_n U) - 1)} \quad (3-26)$$

which can be used to determine the  $\beta_n$  values based upon the initial values  $S_n$  and the upper limit  $U$ .

Now that we know the form of prior estimate of the pdf  $p(\mathbf{m})$ , the next step is to determine the posterior estimate of the pdf  $q(\mathbf{m})$ , by minimizing the relative entropy of  $q(\mathbf{m})$  relative to  $p(\mathbf{m})$ .

$$\text{Minimize } \int q(\mathbf{m}) \ln \left[ \frac{q(\mathbf{m})}{p(\mathbf{m})} \right] d\mathbf{m} \quad (3-27)$$

$$\text{such that } \int q(\mathbf{m}) d\mathbf{m} = 1 \quad (3-28)$$

$$\text{and } \int_M q(\mathbf{m}) \left[ \sum_{n=1}^N g_{jn} m_n \right] d\mathbf{m} = \bar{d}_j ; \quad j = 1, 2, \dots, M \quad (3-29)$$

Again using Lagrangean optimization, the resulting equation for the posterior pdf is (Woodbury and Ulrych 1993).

$$q(\mathbf{m}) = p(\mathbf{m}) \exp \left[ -1 - \mu - \sum_{j=1}^M \lambda_j \sum_{n=1}^N g_{jn} m_n \right] \quad (3-30)$$

The equation can be rewritten by grouping constants, substituting for  $p(\mathbf{m})$ , and applying the unity constraint (3-28) (Woodbury and Ulrych 1993).

$$q(\mathbf{m}) = \prod_{n=1}^N \frac{-a_n}{\exp(-a_n U) - 1} \exp[-m_n a_n] \quad (3-31)$$

where  $a_n$  is defined as

$$a_n = \beta_n + \sum_{j=1}^M \lambda_j g_{jn} \quad (3-32)$$

Equation 3-31 represents the posterior estimate of the multivariate pdf for the set of all model parameters  $\mathbf{m}$ . The posterior estimate for the pdf of each individual model parameter  $m_n$  is



$$q(m_n) = \frac{-a_n}{\exp(-a_n U) - 1} \exp[-m_n a_n] \quad (3-33)$$

Knowing the form of the posterior estimate of the pdf  $q(m_n)$ , we can now solve for the expected value of each model parameter.

$$\bar{m}_n = \int_0^{U_n} m_n q(m_n) dm_n \quad (3-34)$$

where  $\bar{m}_n$  is the expected value for  $m_n$ . By substituting (3-33) into (3-34) and integrating, the expected value is found to be.

$$\bar{m}_n = \frac{\exp(-a_n U) a_n U + \exp(-a_n U) - 1}{a_n (\exp(-a_n U) - 1)} \quad (3-35)$$

Now, substituting (3-34) into (3-19) and solving,

$$\bar{d}_j = \sum_{n=1}^N g_{jn} \bar{m}_n \quad (3-36)$$

Equation (3-36) represents the expected (mean) value of the  $j^{th}$  observation  $\bar{d}_j$  in terms of the transfer function  $g_{jn}$ , and expected values of the unknown model parameters  $\bar{m}_n$ .

Where  $\bar{m}_n$  is a function of the Lagrangean multipliers  $\beta_n$  and  $\lambda_j$ . It should be noted that equation (3-16) is essentially the same as equation (2-1) used in the development of the mathematical model in Chapter 2. The subtle distinction being that (3-36) is written explicitly in terms of the expected values of the unknown model parameters (flux)  $\bar{m}_n$ , and observed concentration  $\bar{d}_j$ .

In order to solve for the parameter expected values the Lagrangean multipliers must be determined. The  $\beta_n$  are estimated based upon the initial values supplied for the unknown model parameters. Then, the  $\lambda_j$  values are estimated through an iterative

procedure that upon convergence results in optimal  $\bar{m}_n$  values and posterior estimates of each pdf  $q(m_n)$ .

A general MRE algorithm is shown in Figure 3-1. Once the algorithm has converged, the final  $a_n$  values (based upon optimal  $\lambda_j$  values) can be used to estimate the model parameter confidence intervals. This is done by calculating the individual cumulative density function (cdf) for each model parameter. The cdf is calculated by integration of the pdf,

$$\int_0^{m'} q(m_n) dm_n = Q(m') \quad (3-37)$$

where  $q(m_n)$  is the posterior estimate of the pdf as defined in equation (3-33),  $m'$  is an arbitrary model parameter value ( $0 \leq m' \leq U$ ), and  $Q(m')$  is the posterior estimate of the cdf. Substituting (3-33) into (3-37) and solving the integral yields,

$$Q(m') = \frac{\exp(-a_n m') - 1}{\exp(-a_n U) - 1} \quad (3-38)$$

which represents the posterior estimate of the cdf in terms of the Lagrangean multipliers ( $a_n$ ), the parameter upper limit  $U$ , and the parameter value  $m'$ . Solving (3-38) for  $m'$  yields the following equation,

$$m' = \frac{\ln[Q(\exp[-a_n U] - 1) + 1]}{a_n} \quad (3-39)$$

which can be used to calculate the model parameter value corresponding to a specified probability level  $Q$ .

#### Algorithm application

In order to implement the minimum relative entropy algorithm with the mathematical model developed in Chapter 2 (Figure 2-1) one must first establish the



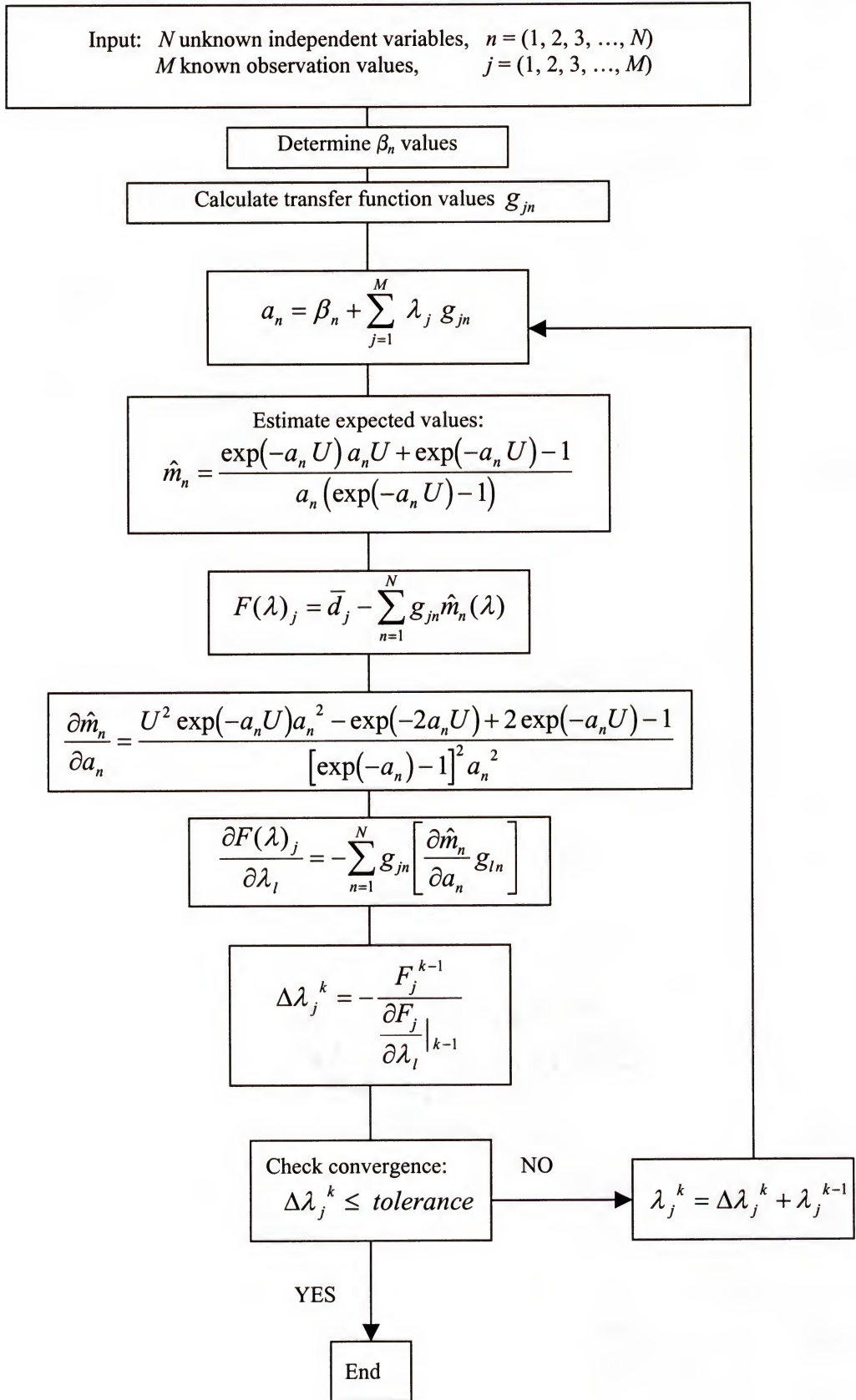


Figure 3-1. General minimum relative entropy algorithm

upper and lower limits  $U$  and  $L$ . As mentioned previously, care must be taken when assigning the upper and lower limits. It is important to select a sufficiently large value of  $U$  that will allow for all expected values of  $m_n$  to occur. The value of  $U$  can always be lowered once it is determined that none of the  $m_n$  values exceed  $U$ . It should also be noted that  $\beta_n$  becomes zero and causes a discontinuity when  $S_n = U/2$ . This should be kept in mind while assigning the initial values for  $S_n$ .

Next, the number of cells within the flux plain  $N$  and the number of observations  $M$  must be established along with the coordinates for the  $N$  flux cells and  $M$  observation locations. Then,  $N$  initial flux values  $S_n$  and the corresponding  $\beta_n$  values must be provided for each flux cell and  $M$  initial  $\lambda_j$  values must be provided for each observation. The initial  $\beta_n$  values can be estimated using the first-order Taylor series approximation to equation (3-26)

$$\beta_n = \frac{3}{2S_n^2} [U - 2S_n] \quad (3-40)$$

However, as noted by Woodbury and Ulrych (1993) this approximation is inadequate when the initial estimate  $S_n$  is close to the upper or lower limits ( $U$  or  $L$ ). The easiest method for determining the initial  $\beta_n$  values is to use equation (3-26) along with the Goal Seek function in Microsoft Excel. The  $\lambda_j$  values can initially be set the same for all observations. Next, any nonlinear parameter estimation algorithm can be used to determine the  $\beta_n$  and  $\lambda_j$  values that best match the observed concentration data. For this study a modified Powell hybrid algorithm was used to determine the  $\beta_n$  and  $\lambda_j$  values which were then used to estimate the flux expected values  $m_n$ .



### Random Search Techniques

Random search techniques are optimization methods that search randomly through a solution space in order to locate the “globally” optimal solution. Unlike gradient-based methods, random search techniques require only objective function values; no derivative calculations are required.

Early random search techniques such as the random walk method were based upon brute force. The algorithm literally searched randomly through all possible solutions until predetermined convergence criteria were met or in the case of combinatorial optimization, until all possible permutations were exhausted. If the algorithm was allowed to search long enough, it would in theory arrive at the optimal solution. But, as a system increased in size and complexity, the necessary time for arriving at an optimal solution also increased.

When first introduced, most random search methods were considered too computationally expensive. However, the increase of computer processing speeds and the implementation of probabilistic decision rules within the random search framework has made random search techniques highly feasible modeling tools. There are a wide variety of random search techniques documented in the literature, simulated annealing, evolutionary algorithms, and neural networks to name a few (Duan et al., 1994; Rogers et al., 1995; Zheng and Wang 1996; Wang and Zheng 1998). Each method has its own merits and pitfalls but the more recent methods all have two similarities: they require only objective function information to determine convergence, so derivative calculations are not required; and they all implement probabilistic transition rules, which allow them to avoid local minima in an effort to move towards a global minimum. The drawbacks of random search techniques are the computational cost and the fact that although

theoretically they should locate a global optimum provided enough time, the final estimates cannot be represented in terms of uncertainty.

### Simulated Annealing

The method of simulated annealing, introduced by Kirkpatrick et al. (1983), is based on an analogy with thermodynamics, specifically with the way that metals cool or anneal. When a metal is heated to its melting point it becomes a liquid. At high temperatures the molecules of a liquid move randomly with respect to one another. If the liquid is cooled slowly, thermal mobility is gradually decreased, and the atoms are often able to line up and form a pure crystal. The crystal represents the optimal state of minimal energy for the system. If a liquid metal is cooled quickly a polycrystalline or amorphous state with slightly higher energy is formed. The slow cooling process is the key to obtaining the state of optimal (minimal) energy.

The annealing process can be simulated as a method for numerical optimization. The algorithm is based on the Boltzmann probability distribution:

$$P[\bar{E} = E(i)] = \frac{1}{Z(T)} \cdot \exp\left(-\frac{E(i)}{kT}\right) \quad (3-41)$$

where  $\bar{E}$  represents the thermal energy of the system (a random variable),  $E(i)$  is the energy corresponding to state  $i$ ,  $k$  is Boltzmann's constant, and  $Z(T)$  is a normalization factor known as the partition function which depends on the system temperature  $T$ . The factor  $\exp\left(-\frac{E(i)}{kT}\right)$  is known as the Boltzmann factor. As the temperature decreases, the Boltzmann probability distribution is weighted towards states with lower energy. When the temperature approaches zero, only the minimal energy states have a non-zero probability of occurrence.



In 1953, Metropolis et al. applied the principles of statistical mechanics to numerical calculations by simulating the evolution to thermal equilibrium of a solid with a fixed temperature (Laarhoven and Aarts, 1987). The general procedure is that given a current state, a simulated thermodynamic system is perturbed. If the difference in energy  $\Delta E$  is negative (which results in state of lower energy) then the process is continued in the new state. If the difference in energy is positive  $\Delta E \geq 0$  (which results in state of higher energy), the probability of acceptance of the new state is determined as,

$$p = \exp\left(-\frac{\Delta E}{kT}\right) \quad (3-42)$$

This process of typically accepting states of lower energy, but not excluding states of higher energy is known as the Metropolis algorithm. By not excluding perturbations that result in states of higher energy, the Metropolis algorithm is capable of backing out of local minima in favor of moving toward a more global minimum.

The method of simulated annealing takes the analogy with thermodynamics one step further. While the Metropolis algorithm simulates thermal equilibrium at a constant temperature, the method of simulated annealing is a series of Metropolis algorithms evaluated through a sequence of decreasing temperatures. The system perturbations are produced so that they are proportional to the current temperature. So, as the temperature decreases, so does the scale of perturbations. As before, if the resulting change in energy  $\Delta E$  is negative, the perturbation is always accepted, and if the change in energy is positive  $\Delta E \geq 0$ , the probability of acceptance is determined by equation (3-42). Thus, at higher temperatures, the algorithm creates larger perturbations, and the probability of accepting a move in the direction of higher energy is greater. As the temperature decreases, the scale of the perturbations decreases, and the probability of accepting



moves in the direction of higher energy is reduced. In terms of optimization, this means that at higher temperatures the algorithm is more active—it searches a larger portion of the solution space and is more capable of moving out of local minima. As the temperature decreases the algorithm activity decreases—the search is refined to smaller and smaller portions of the solution space, and the ability to move out of local minima is reduced.

When applying simulated annealing to optimization problems, the solution space corresponds to the various phases of a material as it cools, the objective function  $f(x)$  is analogous to energy  $E$ , and the analog of temperature  $T$ , is an algorithm control parameter which dictates the scale of perturbations and the probability of accepting uphill moves. The process requires that an annealing schedule be established which controls how and when  $T$  is lowered from high to low values. The procedure for extending the method of simulated annealing to a non-thermodynamic system is summarized in the following four steps (Press et al., 1992):

1. Provide a mathematical description for all possible system configurations.
2. Determine an objective function, which will be minimized.
3. Establish a control parameter  $T$  (analog of temperature) and an annealing schedule which tells how  $T$  is lowered and when.
4. Decide upon a method for producing random changes in the system configuration.

A general simulated annealing algorithm is shown in Figure 3-2. The general simulated annealing procedure described above was initially developed for application to combinatorial optimization. With combinatorial optimization problems, the goal is to

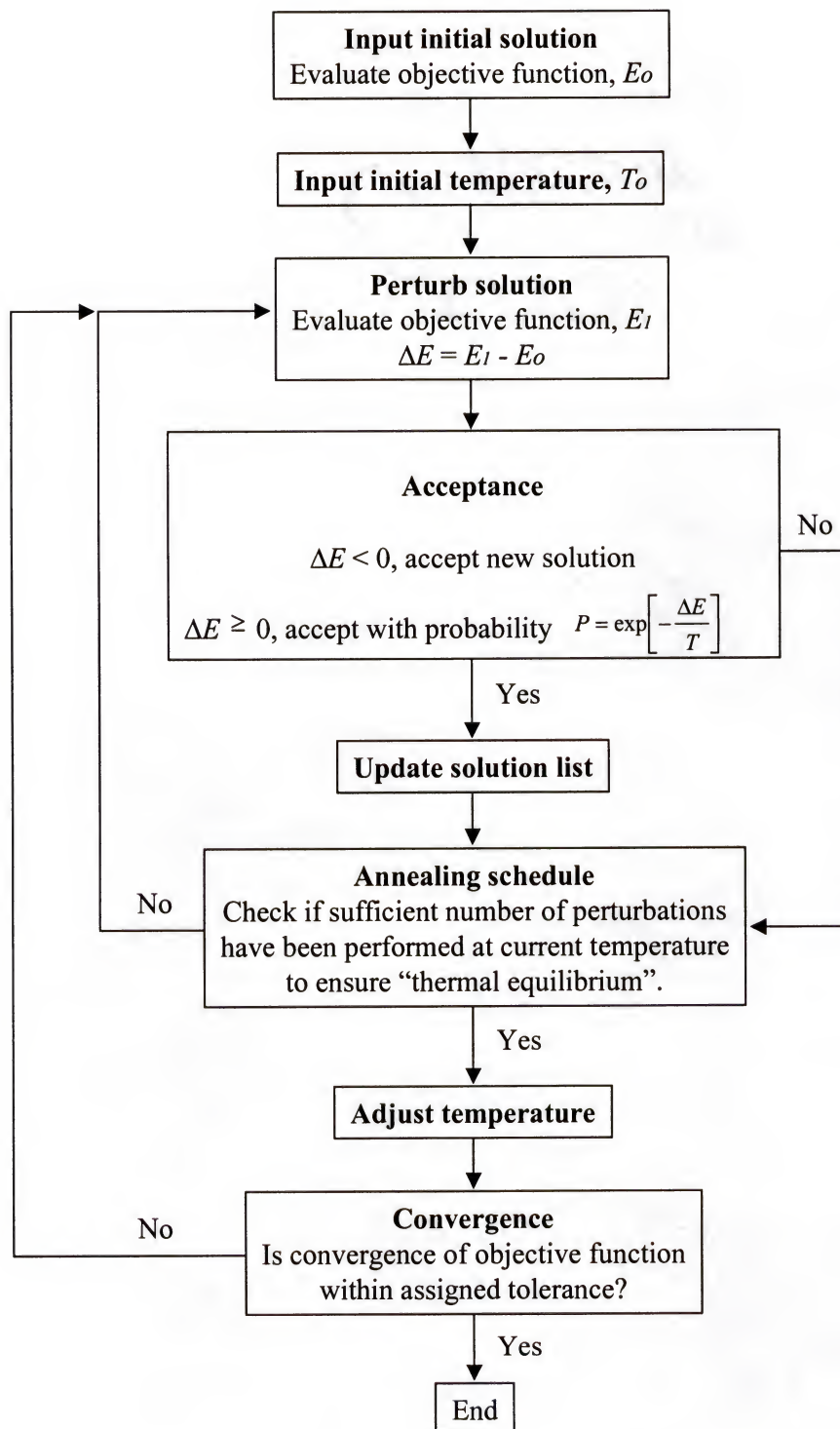


Figure 3-2. General simulated annealing algorithm.

determine the optimal permutation of a finite set of system parameters or control variables. The final solution is usually in the form of a finite list of integers that represent the optimal system configuration. Because the system consists of a finite number of possible configurations, determining a method for introducing random changes in configuration is essentially a matter of shuffling through all possible choices. This is not usually difficult, and there are often several options to choose from. However, with the case of continuous variables, determining a method for producing random changes in system configuration is not as simple.

When dealing with the optimization of  $N$  continuous variables the simulated annealing procedure has to be recast. Now, the goal is to determine the minimum of some function  $f(\mathbf{x})$  where  $\mathbf{x}$  is an  $N$ -dimensional vector representing the state of the system. The procedure for applying the method of simulated annealing to a system of continuous variables is as follows (Press et al., 1992):

1. The state of the system is described by the point  $\mathbf{x}$  (an  $N$ -dimensional vector).
2. The objective function to be minimized is  $f(\mathbf{x})$ .
3. The control parameter  $T$  is as before, an analog to temperature with an annealing schedule by which it is gradually reduced.
4. There must be a method for producing random changes in system configuration by taking random steps from  $\mathbf{x}$  to  $\mathbf{x} + \Delta\mathbf{x}$ .

The final element is the most critical, as the efficiency of the algorithm will be dependent upon the method for generating random changes. Many methods for producing random change exist, but as the number of variables increases most methods become inefficient, especially as convergence to a minimum is approached. For this study, the technique



used to produce random change within the simulated annealing algorithm was a modified downhill simplex method adapted from Press et al. (1992).

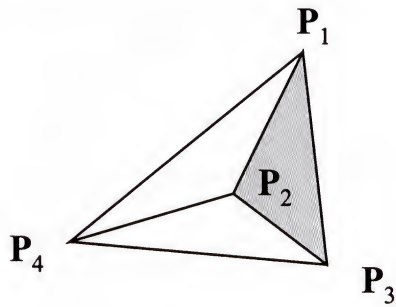
### Downhill simplex method

The downhill simplex method is a random search, nonlinear, multivariate optimization technique that was developed by Nelder and Mead (1965). As with all random search techniques the method requires only function evaluations and no derivative calculations. The premise is based upon the metamorphosis of a geometrical shape called a simplex. In  $N$  dimensions, a simplex is a shape consisting of  $N + 1$  points (or vertices). In two dimensions, a simplex is a triangle, and in three dimensions it is a tetrahedron. Figure 3-3 shows a simplex in three dimensions, and provides a general geometrical description and a mathematical description as applied in optimization. In order to extend the geometric analog of a simplex to optimization, the number of dimensions ( $N$ ) corresponds to the number of independent variables. Each point (or vertex) of the simplex is a set of  $N$  independent variables representing a possible solution to the objective function. Mathematically, the simplex is represented as an array consisting of  $N$  columns and  $N + 1$  rows. Each of the  $N + 1$  rows within the array is a vector of  $N$  independent variables corresponding to a vertex of the simplex and a possible solution to the objective function.

With most methods for multivariate (multi-dimensional), nonlinear minimization the usual way to initiate a search is to provide an initial guess—a point consisting of  $N$  initial values for the independent (control) variables. The optimization algorithm then uses this initial guess as the starting point for its downhill movement through  $N$ -dimensional space in search of a minimum.

Number of dimensions:  $N = 3$   
 $N + 1 = 4$

Geometrical (simplex with  $N + 1$  vertices)



$$\mathbf{P}_i = (x_i, y_i, z_i)$$

$$i = 1, 2, \dots, N + 1$$

Mathematical (simplex = array with  $N$  columns and  $N + 1$  rows)

	$N$			
	$x_1$	$y_1$	$z_1$	$= \mathbf{P}_1$
	$x_2$	$y_2$	$z_2$	$= \mathbf{P}_2$
	$x_3$	$y_3$	$z_3$	$= \mathbf{P}_3$
$N + 1$	$x_4$	$y_4$	$z_4$	$= \mathbf{P}_4$

Figure 3-3. General description of a simplex in three dimensions.

The downhill simplex method does not start with a single point, but with an initial simplex of  $N + 1$  points. Once the initial simplex is established, the downhill simplex method performs a series of steps in order to manipulate and move the simplex through the solution space with the goal of contracting around a global minimum. Movement of the simplex is produced by a combination of three types of steps: reflection, expansion, and contraction. These steps are illustrated in Figure 3-4. All steps are predicated upon the value of the objective function estimated at each of the simplex vertices. A reflection involves moving the highest point on the simplex (corresponding to the largest objective function value) through the opposite face of the simplex to a lower point (corresponding to a point of lesser value for the objective function). When possible, a reflection will be accompanied by an expansion, an attempt to take larger steps in the direction of a possible minimum. When a valley within the solution space is encountered the method contracts the simplex in the transverse direction toward the valley floor and attempts to move through the valley. If a situation arises where the simplex must pass through a very small region within the solution space, the simplex is contracted in all directions toward the current lowest point.

The manipulations discussed above allow the simplex to move through the solution space in search of a minimum value. However, only downhill movements are allowed (in the direction of a lower objective function value). There is no mechanism in the original downhill simplex method to allow for possible uphill movements in order to escape a local minimum. By combining the downhill simplex method with the Metropolis procedure the robust searching abilities of the downhill simplex method are



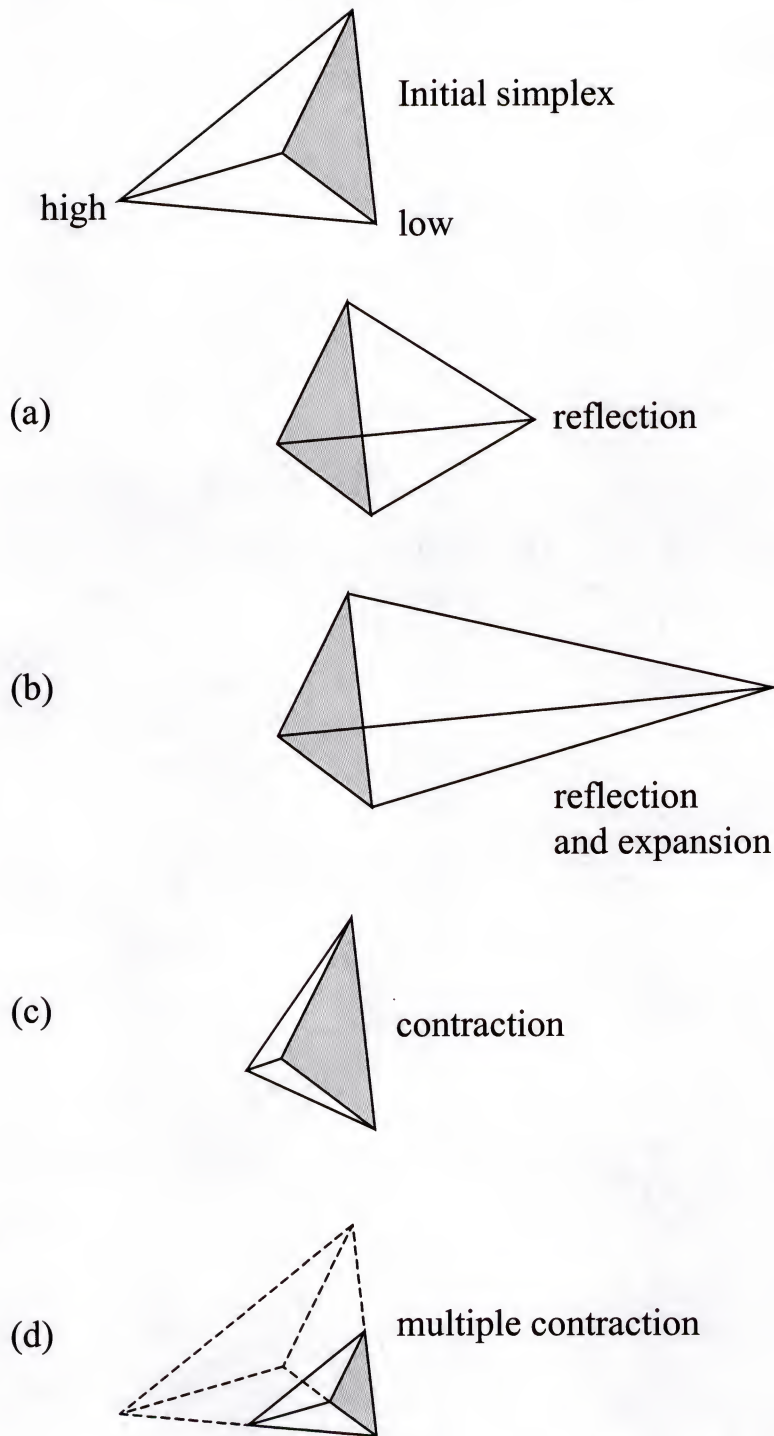


Figure 3-4. Possible outcomes for a step in the downhill simplex method. The simplex at the beginning of the step, here a tetrahedron, is shown, top. The simplex at the end of the step can be any one of (a) a reflection away from the high point, (b) a reflection and expansion away from the high point, (c) a contraction along one dimension from the high point, or (d) a contraction along all dimensions towards the low point. (Adapted from Press et al., 1992).

improved by allowing the simplex to move uphill away from local minima in hopes of moving toward a global minimum.

### Simulated annealing with modified downhill simplex

As stated previously, the critical element in implementing simulated annealing with continuous variables is the method for generating random steps from  $\mathbf{x}$  to  $\mathbf{x} + \Delta\mathbf{x}$ . Combining the search techniques of the downhill simplex method with the thermodynamic analog of the Metropolis algorithm provides a unique process for generating random changes. In order to incorporate the downhill simplex method within a simulated annealing algorithm, the point  $\mathbf{x}$  is replaced by a simplex of  $N + 1$  points. The steps used to manipulate the simplex are the same as discussed previously: reflection, expansion, and contraction.

The Metropolis procedure is implemented indirectly by performing a random thermal fluctuation on each vertex of the simplex prior to each simplex manipulation. The thermal fluctuation is produced by adding a positive, logarithmically distributed random variable, proportional to the temperature  $T$ , to the objective function value associated with each vertex of the simplex. This is done using the following general relationship:

$$f(P_i)' = f(P_i) + T \cdot \text{ran} \quad (3-43)$$

Where  $P_i$  represents vertex  $i$  of the simplex,  $f(P_i)$  is the objective function evaluated at  $P_i$ ,  $T$  is the current annealing temperature,  $\text{ran}$  is a logarithmically distributed random variable, and  $f(P_i)'$  is the perturbed objective function value corresponding to vertex  $i$ . Then, it is the perturbed objective function values that are used to determine which vertex



of the simplex should be replaced. Next, the same type of random variable is subtracted from the function values for each of the new points that are considered as replacements.

Because the objective function values are randomly perturbed, the order of the perturbed values is not necessarily proportional to the order of the actual function values. This means that the lowest objective function value may not correspond to lowest perturbed value, and the highest objective function value may not correspond to the highest perturbed value. Consequently, a simplex reflection from the highest perturbed value to a lower perturbed value may actually result in a reflection from a low objective function value to a higher objective function value (i.e. an uphill move). Even though the probability of accepting an uphill move is not explicitly determined using the Boltzmann factor (3-42), the Metropolis procedure is maintained because a mechanism for random acceptance of uphill moves does exist. Also, because the perturbations are proportional to the temperature, the variable activity of the search is maintained—as the temperature decreases the scale of perturbations and the probability of accepting uphill moves decrease.

### Annealing schedule

The rate at which the system temperature is lowered is controlled by the annealing schedule. As implied by the name, the schedule represents a strategy for how much and how often the temperature is decreased. The success of a simulated annealing algorithm is primarily dependent upon how slowly the system is cooled. Because the activity of the algorithm is dependent upon the temperature, if the temperature is decreased too quickly the search is inefficient—the algorithm does not explore a large enough portion of the solution space and there is very little chance of encountering a global minimum. Conversely, if the temperature is lowered excessively slow the algorithm becomes



inefficient based upon time constraints (Although, with current computer processing speeds this is not always an issue). There are numerous strategies for annealing schedules from very simple to extremely complex. The simplest type of schedule being a procedure, such that after every  $n$  moves the temperature is decreased by a predetermined value  $\Delta T$ . Each problem represents its own intricacies and for this reason, the choice of annealing schedule is usually made on a case-by-case basis.

#### Algorithm application

In order to implement the simulated annealing algorithm with the mathematical model developed in Chapter 2 (Figure 2-1) the number of dimensions (cells within the flux plain)  $N$ , the number of points within the initial simplex  $N+1$ , and the number of observations  $M$ , must be established. Then the coordinates of the flux cells and observation locations must be provided along with the initial flux values and observation data. The final input parameter is the annealing temperature and schedule. For the algorithm applied in this study, the annealing schedule is input as a list of starting temperatures and the corresponding maximum number of iterations to be performed at each temperature. The amount of input required to run the simulated annealing algorithm is minimal, but time must be spent to familiarize one self with the system-specific efficiency of the annealing schedule. An annealing schedule that works well for one system may not be as efficient for another. However, once an annealing schedule is established that works well for a given system, it should work well for most cases (realizations) within that system.

#### Shuffled Complex Evolution

Shuffled complex evolution (SCE) represents an evolutionary strategy for global optimization. First presented by Duan et al. (1992), SCE combines the methods of four

optimization techniques: the simplex procedure of Nelder and Mead (1965), the concepts of controlled random search (Price 1987), competitive evolution (Holland 1975), and the process of complex shuffling (Duan et al., 1992; Duan et al., 1993). Like all random search techniques, the method uses only objective function values (no derivative calculations are necessary) and probabilistic decision rules are used to help avoid becoming trapped in local minima.

Because SCE is an evolutionary strategy the method conducts a search from one population of solutions to another, not from individual point to point. The general SCE procedure begins by randomly selecting an initial population of points from the solution space. The initial population is then divided into complexes (or communities) and each complex is made to evolve. Evolution is induced by utilizing a form of simplex methodology to randomly perturb the members of each complex while directing evolution toward locations of improved (decreased) objective function values. Periodically throughout the evolution process the entire population is shuffled by randomly exchanging members between complexes. Given a sufficiently large initial population, the entire population should converge toward the global minimum. The key point being that the initial population has to be sufficiently large. Determining what is sufficiently large is somewhat problem specific. Duan et al., (1994) provide guidelines for establishing appropriate algorithm parameters that control population size, but first the algorithm must be discussed in more detail. A general SCE algorithm is shown in Figure 3-5, and the procedure is presented below.

Given a system consisting of  $n$  unknown parameters ( $n$ -dimensions), let  $\mathbf{x}$  be a point representing one possible solution. The point  $\mathbf{x}$  is an  $n$ -dimensional vector



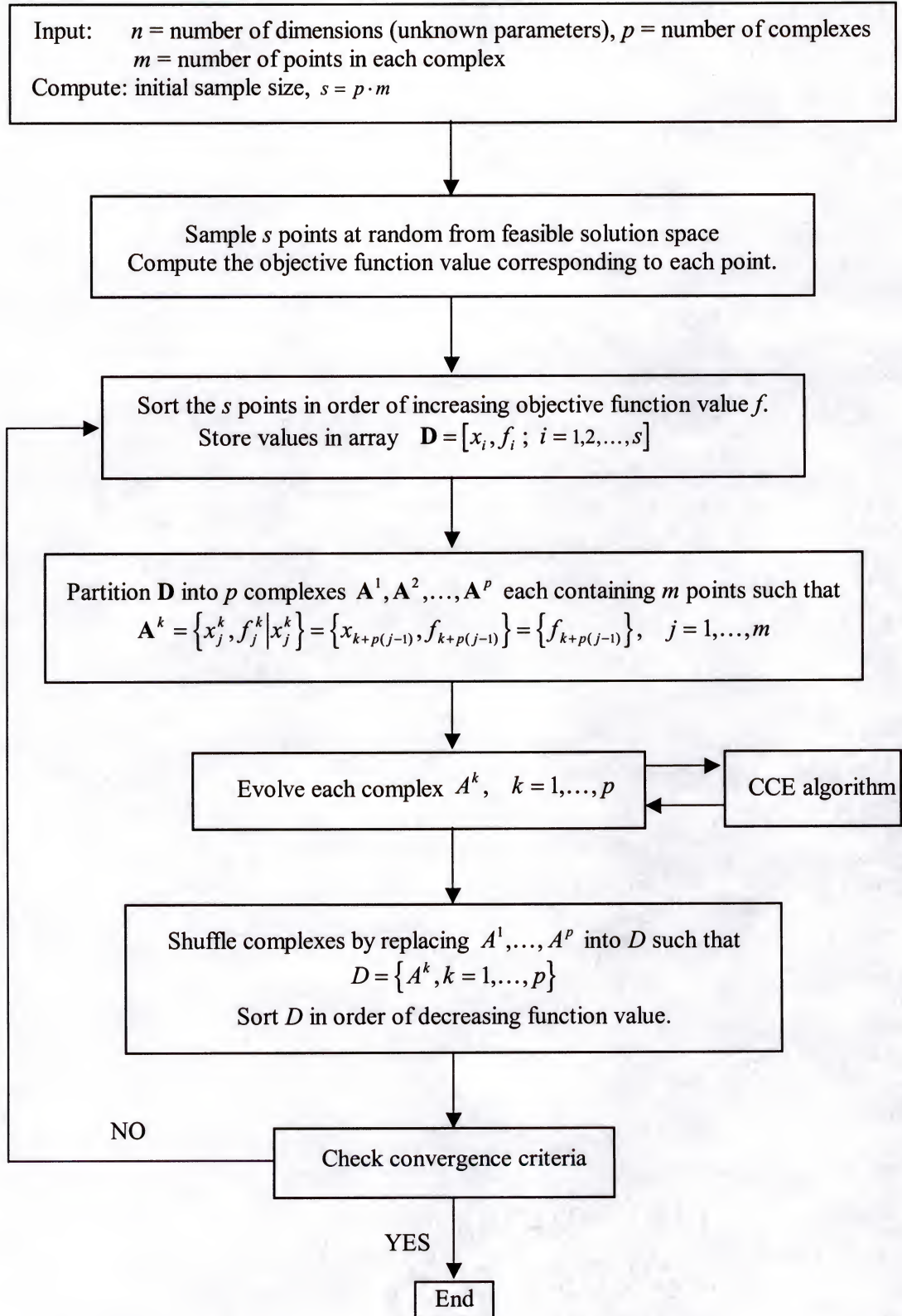


Figure 3-5: A general SCE algorithm (adapted from Duan et al., 1992).



$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad (3-44)$$

where each element of  $\mathbf{x}$  represents a feasible value for one of the  $n$  unknown parameters.

Let the objective function value evaluated at  $\mathbf{x}$ ,  $f(\mathbf{x})$  be represented as  $f$ .

The general SCE procedure requires that an initial population of  $s$  points be randomly selected from the solution space, and then subdivided into smaller communities (complexes). In order to do this we must first define  $p$  is the number of complexes and  $m$  is the number of points in each complex. Both  $p$  and  $m$  are parameters that must be selected prior to application. Based upon the selected values, the initial population or sample size can be defined as

$$s = pm \quad (3-45)$$

Once the  $s$  points are selected randomly from the feasible solution space, the points and their corresponding objective function values can be stored in an array  $\mathbf{D}$

$$\mathbf{D} = \begin{bmatrix} \mathbf{x}_1, f_1 \\ \mathbf{x}_2, f_2 \\ \vdots \\ \mathbf{x}_s, f_s \end{bmatrix} \quad (3-46)$$

Note that each  $\mathbf{x}_i$  represents an  $n$ -dimensional vector and  $f(\mathbf{x}_i)$  represents a scalar value.

The next step is to sort the  $s$  points within  $\mathbf{D}$  in order of increasing objective function values so that the first row in  $\mathbf{D}$ , ( $i = 1$ ), corresponds to the best (smallest) objective function value. Once the population has been sorted, it is then divided into  $p$  complexes (or communities) each containing  $m$  sets of parameters and corresponding objective function values.

$$\mathbf{A}^k = \begin{bmatrix} \mathbf{x}_1^k, f_1^k \\ \mathbf{x}_2^k, f_2^k \\ \vdots \\ \mathbf{x}_m^k, f_m^k \end{bmatrix} \quad \text{for } k = 1, 2, \dots, p \quad (3-47)$$

where  $\mathbf{A}^k$  represents one of  $p$  complexes all contained within the array  $\mathbf{D}$ . In order to reduce the amount of data storage and to make matrix manipulation more efficient, each row of the array  $\mathbf{D}$  can be defined as an independent vector

$$\mathbf{g}_i = [\mathbf{x}_i, f_i] \quad \text{for } i = 1, 2, \dots, s \quad (3-48)$$

and each row of the complexes  $\mathbf{A}^k$  can be defined as

$$\mathbf{g}_{j,k} = [\mathbf{x}_j^k, f_j^k] \quad \text{for } j = 1, 2, \dots, m \text{ and } k = 1, 2, \dots, p \quad (3-49)$$

Then the complexes can be referenced using an adjusted index system.

$$\mathbf{g}_{j,k} = \mathbf{g}_i \quad \text{where } i = k + p(j-1) \quad (3-49)$$

Figure 3-6 shows a general  $\mathbf{D}$  array containing complexes  $\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^p$ , where each complex contains  $m$  sets of parameters and objective function values.

Once the population has been partitioned into complexes, each complex is evolved using the competitive complex evolution (CCE) algorithm (discussed in the next section). The evolved complexes ( $\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^p$ ) are then placed back into  $\mathbf{D}$ , and the entire population is sorted in order of increasing objective function value. By sorting the entire  $\mathbf{D}$  array the  $p$  independent complexes are disassociated back into the larger population ( $s = pm$ ). This induces a sharing of information between complexes before they are repartitioned and evolved again.

After the population is sorted, convergence is evaluated by determining if the smallest objective function meets a predetermined convergence criterion. If the



$$\mathbf{D} = \left\{ \begin{array}{l} \begin{array}{l} g_1 \left[ \begin{array}{c} x_{1,1}, x_{1,2}, x_{1,3}, \dots \\ x_{2,1}, x_{2,2}, x_{2,3}, \dots \\ \vdots \\ x_{m,1}, x_{m,2}, x_{m,3}, \dots \end{array} \quad \begin{array}{c} , x_{1,n}, f_1 \\ , x_{2,n}, f_2 \\ \\ , x_{m,n}, f_m \end{array} \end{array} \\ \begin{array}{l} g_{m+1} \left[ \begin{array}{c} x_{m+1,1}, x_{m+1,2}, x_{m+1,3}, \dots \\ x_{m+2,1}, x_{m+2,2}, x_{m+2,3}, \dots \\ \vdots \\ x_{2m,1}, x_{2m,2}, x_{2m,3}, \dots \end{array} \quad \begin{array}{c} , x_{m+1,n}, f_{m+1} \\ , x_{m+2,n}, f_{m+2} \\ \\ , x_{2m,n}, f_{2m} \end{array} \end{array} \\ \vdots \\ \begin{array}{l} g_{pm} \left[ \begin{array}{c} \ddots \\ \ddots \\ \ddots \\ \ddots \\ x_{pm,1}, x_{pm,2}, x_{pm,3}, \dots \end{array} \quad \begin{array}{c} \\ \\ \\ \\ , x_{pm,n}, f_{pm} \end{array} \end{array} \end{array} \right\} \begin{array}{l} \mathbf{A}^1 \\ \mathbf{A}^2 \\ \mathbf{A}^p \end{array}$$

Where the population sample size is defined as

$$s = pm$$

and each complex contains  $m$  sets of parameters and objective function values.

Figure 3-6. A general description of the array  $\mathbf{D}$  containing complexes  $\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^p$ .



convergence criteria is not met the algorithm checks to see if the number of complexes can be reduced, thus reducing the computational burden. If the minimum number of complexes  $p_{min}$  is less than the current number of complexes  $p$ , then the complex with the lowest ranking is removed. The SCE procedure continues until the convergence criterion is satisfied or the preset maximum number of iterations is reached.

### Competitive complex evolution

The competitive complex evolution (CCE) algorithm is the primary component of the shuffled complex evolution method. The CCE algorithm uses the downhill simplex method of Nelder and Mead (1965) to randomly generate population offspring. A general CCE algorithm is shown in Figure 3-7 and is outlined below based upon the work of Duan et al., (1992, 1994).

To initialize the CCE algorithm the values of  $q$ ,  $\alpha$ , and  $\beta$  must be assigned. Where  $q$  is the number of points in a sub-complex,  $\alpha$  is the number of consecutive offspring that may be generated by the same sub-complex, and  $\beta$  is the number of evolution steps that each complex will take before complexes are shuffled.

Then, by assuming a triangular probability distribution, assign probability values to each of the  $m$  points in each complex  $A^k$

$$\rho_i = \frac{2(m+1-i)}{m(m+1)}, \quad i = 1, \dots, m \quad (3-50)$$

The triangular probability distribution is used in order to assure that the point with the best corresponding objective function value has the highest probability of being chosen for a sub-complex.

Next, using the assigned probabilities  $q$  points are randomly selected from the complex and placed in a sub-complex. Following the evolution analog, this is the

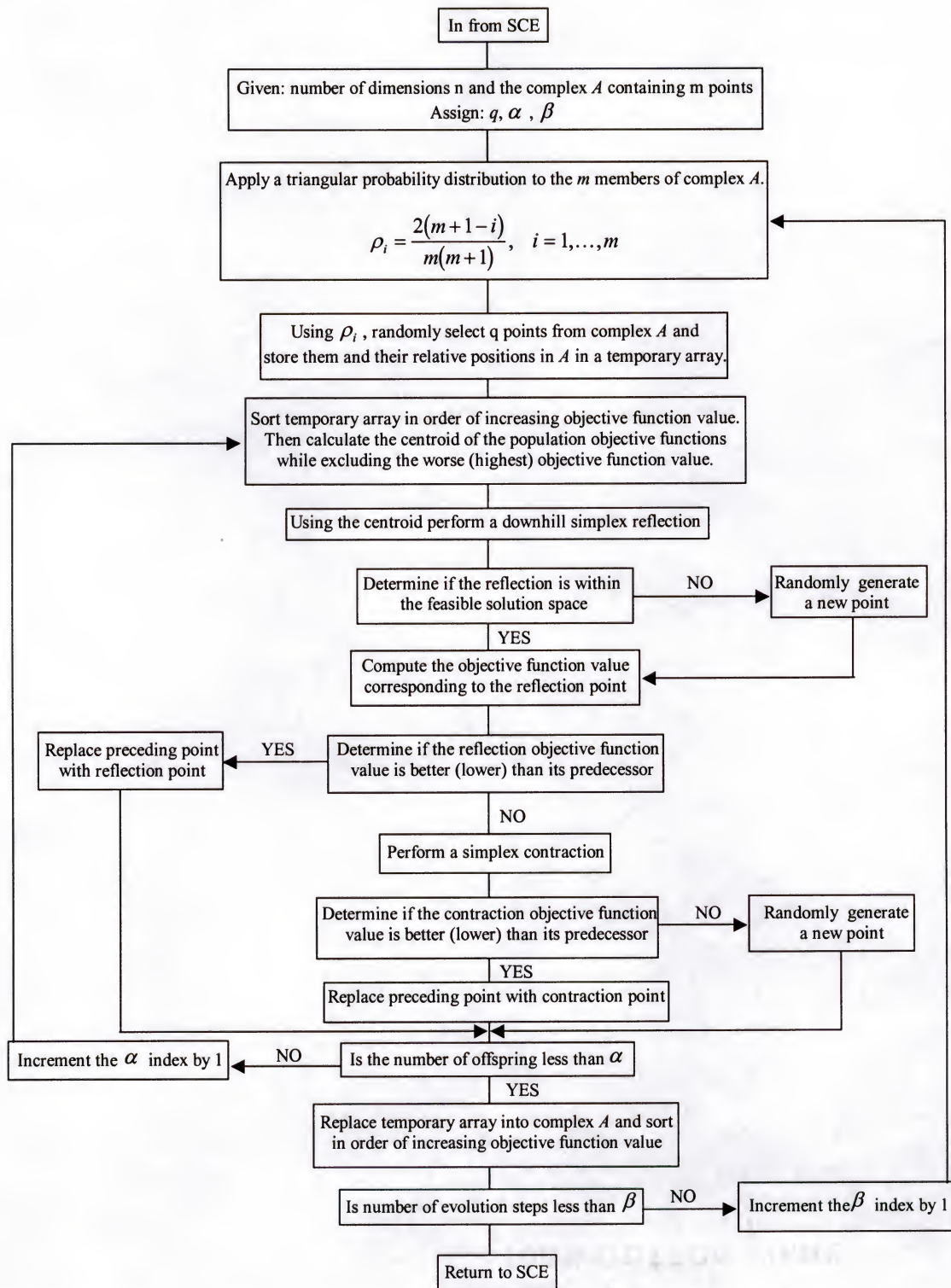


Figure 3-7: A general CCE algorithm (Adapted from Duan et al., 1992).



equivalent of randomly selecting  $q$  parents from the main population (complex) and placing them in a smaller community (sub-complex) where they are allowed to reproduce.

Reproduction is simply the evolutionary analog for producing random change in the set of possible solutions. In this case the set of possible solutions is defined as a complex. In the case of simulated annealing the set of possible solutions is defined as a simplex. They are essentially the same thing—an array containing a set of possible solutions that satisfy the objective function. As with the simulated annealing procedure, the mechanism for producing random change in the shuffled complex evolution procedure is the downhill simplex method of Nelder and Mead (1965). Through a series of reflections and contractions, the downhill simplex method produces new members within the population (complex) of possible solutions by randomly altering the selected parent solutions. The simplex processes of reflection and contraction were presented in the simulated annealing section of this chapter and outlined in Figure 3-4. It should be noted that although shuffled complex evolution and simulated annealing both use the downhill simplex method as their mechanism for random change, how the random changes are applied within the search algorithm is considerably different. Simulated annealing uses the downhill simplex method as a mechanism to randomly move through the set of feasible solutions—it essentially bounces throughout the solution space in search of the optimal solution. Shuffled complex evolution applies the simplex method within a more orderly search process of evolution by sorting and shuffling the array of possible solutions.



### Algorithm application

In order to implement the shuffled complex evolution algorithm with the mathematical model developed in Chapter 2 (Figure 2-1) the number of cells within the flux plain  $n$ , must be established. Then the coordinates of the flux cells and observation locations must be provided along with the initial flux values and observation data.

The success of the SCE algorithm is dependent upon several algorithm parameters. The size of the initial population,  $m$ , being the most important. If the initial population is sufficiently large, the entire population should converge toward a global minimum. Duan et al. (1994) provide guidelines for establishing initial values for the algorithm parameters that affect the population size. Based upon the results of their work, they suggest that the number of sets  $(\mathbf{x}_j, f_j; j = 1, 2, \dots, m)$  within each complex should be determined as  $m = 2n + 1$ .

The other algorithm parameters that must be established are:  $q$  the number of points in a sub-complex;  $p$  the number of complexes,  $p_{min}$ , the minimum number of complexes required in the population;  $\alpha$  the number of consecutive offspring that may be generated by the same sub-complex; and  $\beta$ , the number of evolution steps taken by each complex. Selection of these parameters is critical to the success of the SCE algorithm and often times the best values are difficult to determine. Duan et al. (1994) recommend the following general values:

$$\begin{aligned}
 m &= 2n + 1 \\
 q &= n + 1 \\
 \beta &= m \\
 \alpha &= 1 \\
 p &= p_{min}
 \end{aligned}
 \tag{3-51}$$

Application of these values reduces the burden of parameter selection to only  $p$ —the number of complexes (Agyei, 1997). The value of  $p$  is problem specific, and depends upon the degree of difficulty of the problem. As the number of dimensions (unknown parameters) and difficulty of the objective function increases so does the number of complexes,  $p$ , required to determine the global optimum.



## CHAPTER 4 PHYSICAL MODEL

Three-dimensional physical modeling provides a means to examine complex processes under controlled laboratory conditions. A particular strength of physical modeling is the ability to reproduce experiments; this allows systematic variation in experimental parameters not possible during field experiments. The ability to reproduce experimental conditions also allows verification of experimental results. The physical experiments in this study were designed to provide a sequential set of data to test the capabilities of the mathematical model and solution techniques presented in Chapters 2 and 3.

### Three-Dimensional Aquifer Model

The aquifer model used in this study was constructed at the Air Force Research Laboratory, Tyndall Air Force Base, Florida (Figures 4-1 and 4-2). The system was designed for use with chlorinated solvents; all wetted surfaces are stainless steel or glass, minimizing the partitioning of hydrophobic chemicals to system components. All liquid effluent and vapor streams passed through physical and/or chemical traps, eliminating exposure to hazardous chemicals and allowing quantitative mass balance determinations. In the experiments performed for this study, the model was configured to simulate a homogeneous surficial aquifer. The porous medium was a clean, medium grained silica sand (Flint Silica #14).



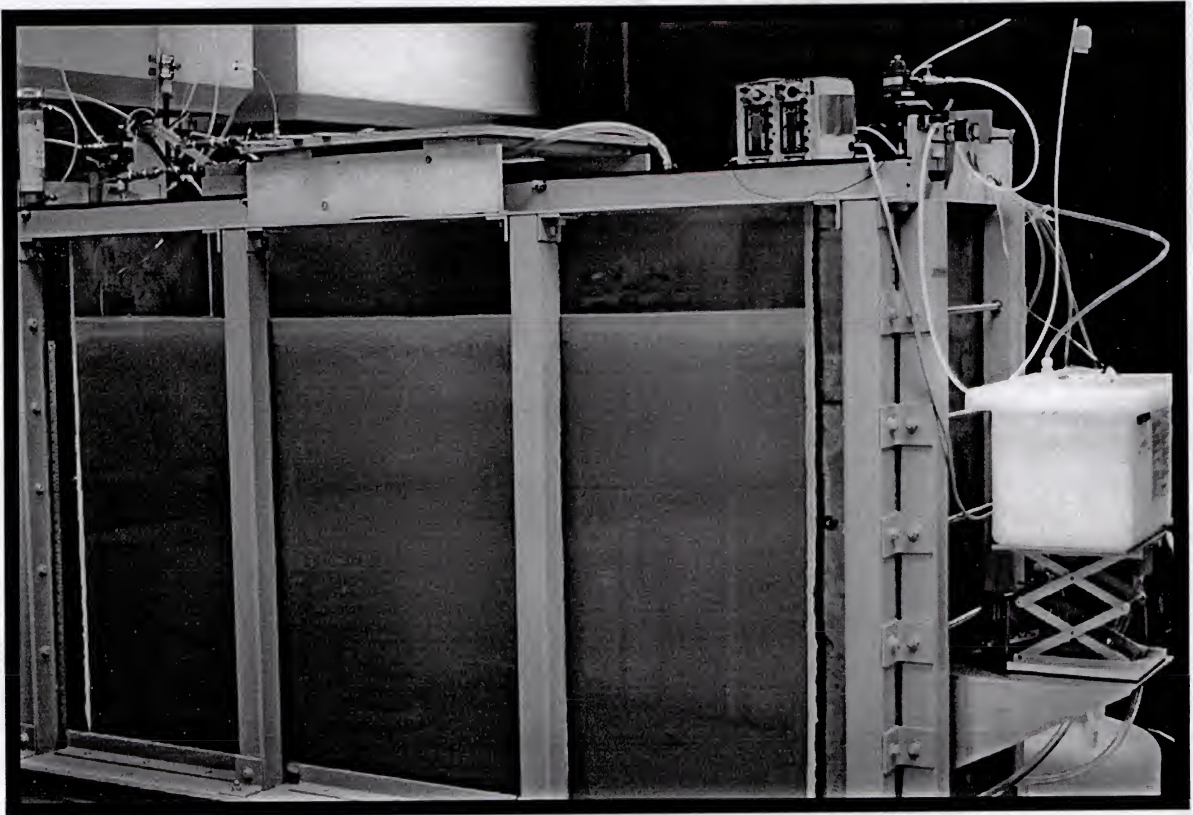


Figure 4-1. Front view of three-dimensional aquifer model (Air Force Research Laboratory, Tyndall Air Force Base, Florida)

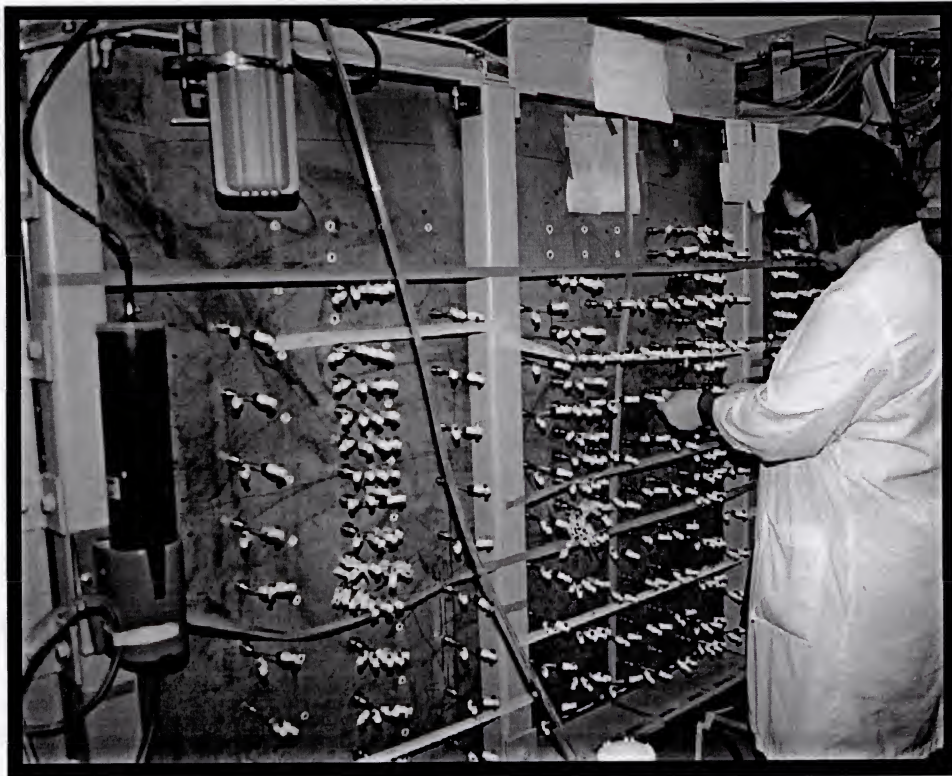


Figure 4-2. Rear view (sampling side) of three-dimensional aquifer model (Air Force Research Laboratory, Tyndall Air Force Base, Florida)



The volumetric flowrate and corresponding pore water velocity were controlled by manipulating the elevations of the inlet and outlet head tanks. The physical, hydrodynamic, and transport characteristics of the system are presented in Table 4-1. The methods for determining these values are discussed in the following sections.

Multi-point samplers. A series of multi-point samplers were emplaced within the flow system providing a three-dimensional distribution of over 500 sampling points within the porous media. The longitudinal distribution (x-z plane) of multi-point samplers is shown in Figure 4-3. The lateral distribution (y-z plane) of sample points varies with location along the longitudinal (x) axis. Some transects have a minimum of 20 sample points while the more extensive transects have up to 72 sample points. Figure 4-4 shows the most extensive sampling transect at  $x = 50$  cm.

The multi-point samplers were designed to allow for measurement of both liquid and vapor phases. The samplers were constructed using 1.3 mm I.D. stainless steel chromatography tubing cut to specific sampling lengths (ranging from 5 to 45 cm). The sampling tubes were joined using a lead free tin solder with an all-purpose soldering flux. The samplers were inserted horizontally through the back wall of the flow system and sealed in place using a stainless steel sheath and Swagelock<sup>®</sup> compression fitting. The extraction ends of the samplers were completed with Teflon syringe sampling ports. The multi-point samplers were constructed in 2-point, 3-point, 4-point, 5-point, and 10-point configurations. Figure 4-5 shows a 2-point sampler assembly.

Placement of porous media and multi-point samplers. The porous media was emplaced dry in a series of 5-cm lifts to a final depth of 1 m. As each lift was completed the multi-point samplers were inserted horizontally and sealed in place. Once the sand



Table 4-1. Physical, hydraulic, and transport properties of the three-dimensional aquifer model.

Parameter	Value
Porous Media	Flint Silica #14 (U.S. Silica, Ottawa, IL)
Porous Media Dimensions (length x width x height; m)	2.0 x 0.5 x 1.0
Median Grain Size ( $d_{50}$ ; mm)	1.2
Porosity ( $n$ )	0.3
Longitudinal Dispersivity ( $\alpha_L$ ; m)	0.001 – 0.003
Transverse Dispersivity ( $\alpha_T$ ; m)	0.0002

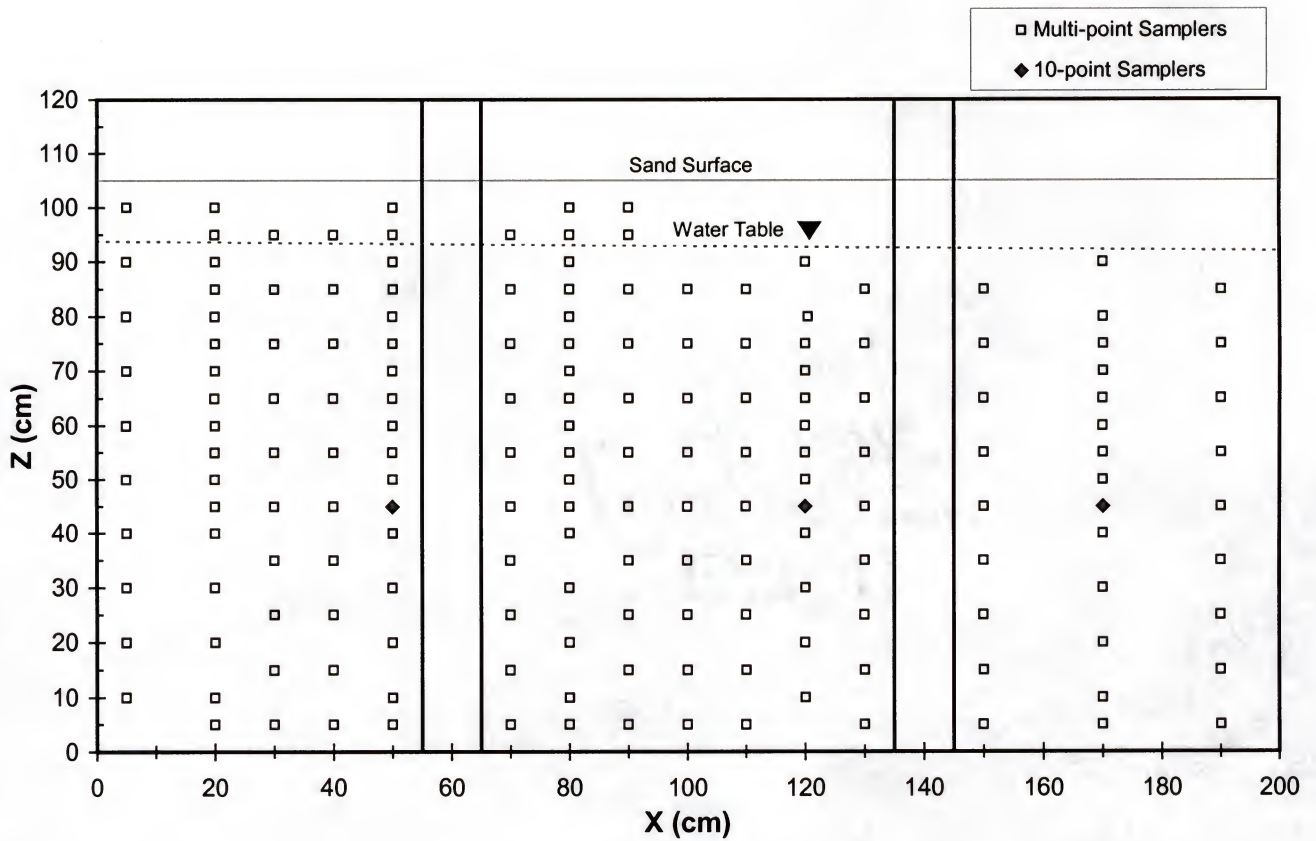


Figure 4-3. Schematic of the multi-sampler distribution along a longitudinal transect of the aquifer model.



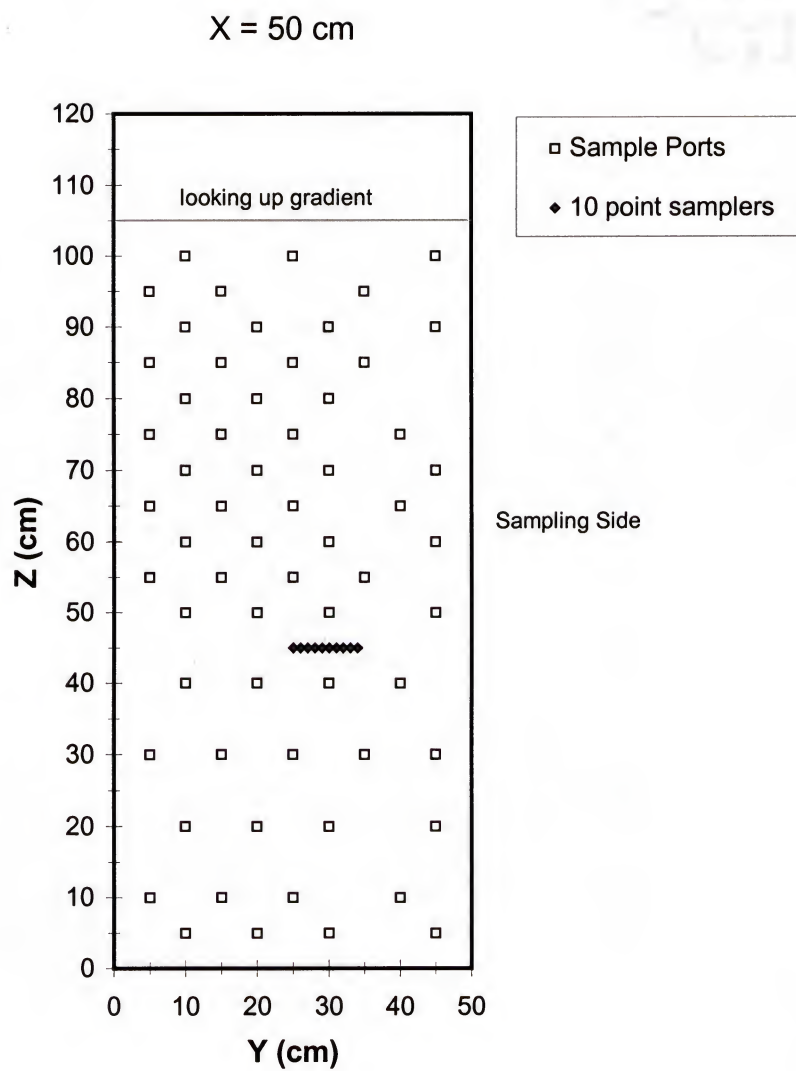


Figure 4-4. Schematic of the sample port distribution along a lateral transect ( $x = 50 \text{ cm}$ ) of the aquifer model.

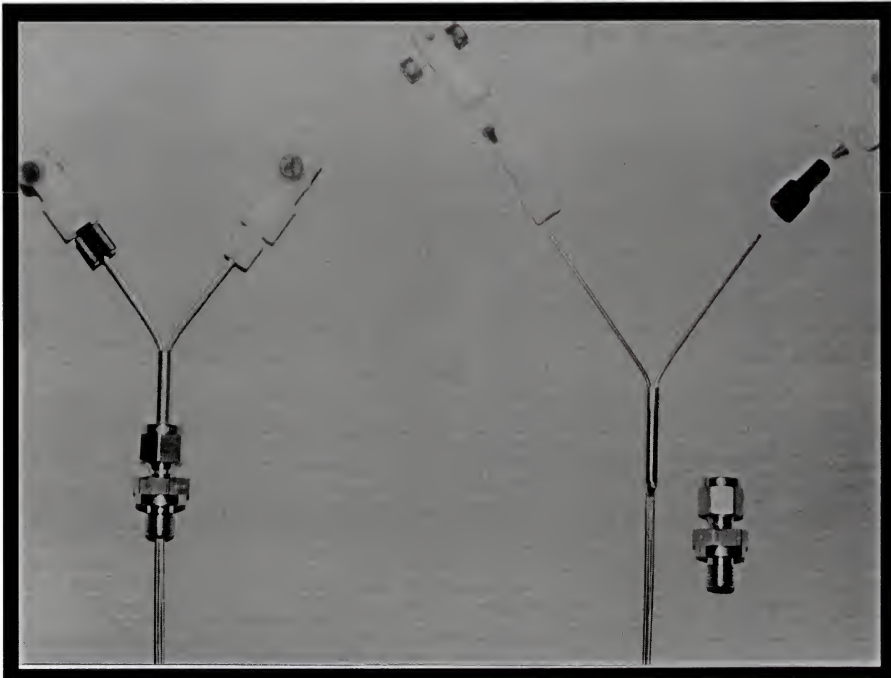


Figure 4-5. Two-point sampler assembly.

and multi-point samplers were in place, water was slowly introduced into the system in a series of 10-cm stages. After the up-gradient head stabilized at each 10-cm increment, flow through the system was stopped and the head was allowed to stabilize for 12 hours before the next stage was introduced. This process was continued until an up-gradient head of 80 cm was established. Flow was then held constant through the system for a period of 7 days.

#### Physical and Hydrodynamic Characterization of the System

The porosity of the system was estimated by packing three columns with porous media and determining their dry mass ( $M_D$ ). The columns were then saturated with water, and their saturated mass ( $M_S$ ) was determined. An average value for porosity was computed as

$$n = \frac{M_w}{\rho_w V_T} \quad (4-1)$$

where  $n$  is the porosity of a given column,  $M_w$  is the water mass ( $M_S - M_D$ ),  $V_T$  is the column volume, and  $\rho_w$  is the density of water ( $1.0 \text{ g/cm}^3$ ). The average porosity was determined to be 0.3.

Given that the model aquifer is of finite extent (2.0 m x 0.5 m x 1.0 m), application of the Dupuit assumptions (Bear 1979) for one-dimensional flow through a homogeneous unconfined aquifer allow for groundwater flow through the system to be estimated using a modified form of Darcy's Equation

$$Q_x = KWh_s \frac{(h_1 - h_2)}{L} \quad (4-2)$$

where  $Q_x$  [ $\text{L}^3\text{T}^{-1}$ ] is the volumetric water discharge through the system,  $K$  [ $\text{LT}^{-1}$ ] is the hydraulic conductivity,  $L$  and  $W$  are the length and width of the model aquifer;  $h_1$  and  $h_2$ ,



are the up- and down-gradient head elevations; and  $h_s$  and  $h_c$  are the saturated and capillary fringe thickness at  $L/2$  (Figure 4-6). Application of the Dupuit assumptions implies that all flow is horizontal, and in this case it is assumed that the water table is essentially linear. By assuming that the water table is linear between the up- and down-gradient heads ( $h_1$  and  $h_2$ ), the saturated thickness  $h_s$  can be estimated using the following relationship.

$$h_s = \frac{h_1 + h_2}{2} + h_c \quad (4-3)$$

Assuming uniform horizontal flow, and that the volumetric flowrate through the system ( $Q_x$ ) can be approximated as the observed mean effluent flowrate, the bulk specific discharge  $q_x$  can be estimated as

$$q_x = \frac{Q_x}{Wh_s} \quad (4-4)$$

and the mean pore water velocity can be estimated as

$$v_x = \frac{Q_x}{nWh_s} \quad (4-5)$$

where  $n$  is the system porosity.

If necessary, equation (4-2) can be used to estimate a bulk hydraulic conductivity value ( $K$ ) for the system. However, it should be noted that because hydraulic conductivity is actually a constant of proportionality in equation (4-2), the estimated  $K$  value is specific to the current up- and down-gradient head values ( $h_1$  and  $h_2$ ) and corresponding saturated thickness ( $h_s$ ).

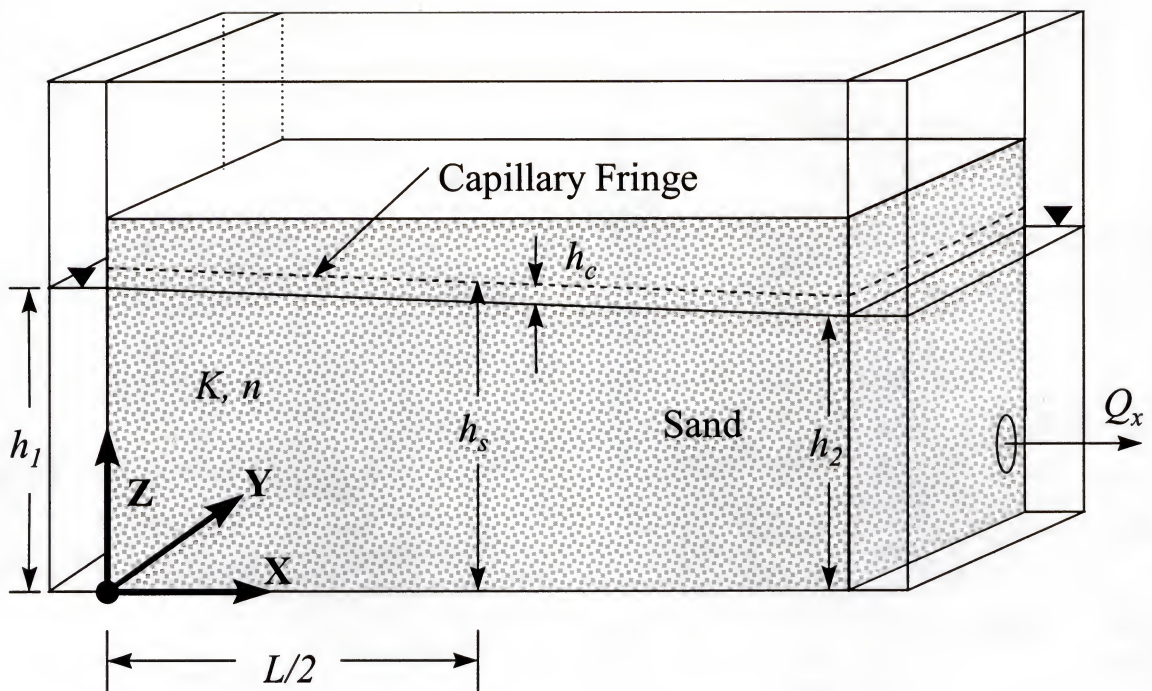


Figure 4-6: Steady unconfined flow through the three-dimensional aquifer model, assuming an essentially linear water table and one-dimensional horizontal flow in the x-direction.

### Transport characterization experiment

In order to determine the longitudinal and transverse dispersivities ( $\alpha_L$  and  $\alpha_T$ ) of the model aquifer, an ionic tracer experiment was performed. The experimental conditions and parameters are summarized in Table 4-2 and a conceptual diagram is shown in Figure 4-7. First, a steady flow field was established with an average saturated thickness ( $h_s$ ) of 0.969 m, and an average pore water velocity of 0.42 m/day. Then a 20-mg/L solution of non-reactive chloride ( $\text{Cl}^-$ ) tracer was injected from a continuous point source located near the inlet boundary at the center of the flow field ( $x = 0.025$  m,  $y = 0.25$  m, and  $z = 0.45$  m). Transient, chloride concentrations were measured at three locations ( $x = 0.5$  m,  $x = 1.2$  m, and  $x = 1.7$  m) along the centerline of the plume (Figure 4-8). Once the plume reached steady state, transverse (horizontal) concentration distributions were measured at the same  $x$ -locations, on three consecutive days (Figure 4-9).

The transverse concentration distributions were measured using 10-point samplers that were designed specifically for this experiment. The 10-point samplers were constructed and installed in order to provide a distribution of sample points with a 1-cm interval ranging from the centerline to the lateral extent of the plume. This allowed for the measurement of steady state transverse concentration profiles within the plume as shown in Figure 4-9. Assuming that the plume was symmetric (implying that the vertical and horizontal transverse dispersion was the same), the observed concentration distributions were used to estimate the transverse dispersivity within the system.



Table 4-2. System parameters for the transport characterization experiment.

Parameter	Value
Chloride Source	
Location (x, y, z; m)	.025, 0.50, 0.45
Concentration (mg/L)	20
Injection Rate ( $Q_s$ ; m <sup>3</sup> /day)	$5.774 \times 10^{-4}$
Saturated Thickness at L/2 ( $h_s$ ; m)	0.969
Average Effluent Water Flowrate ( $Q_x$ ; m <sup>3</sup> /day)	0.061
Average Pore Water Velocity ( $v$ ; m/day)	0.42

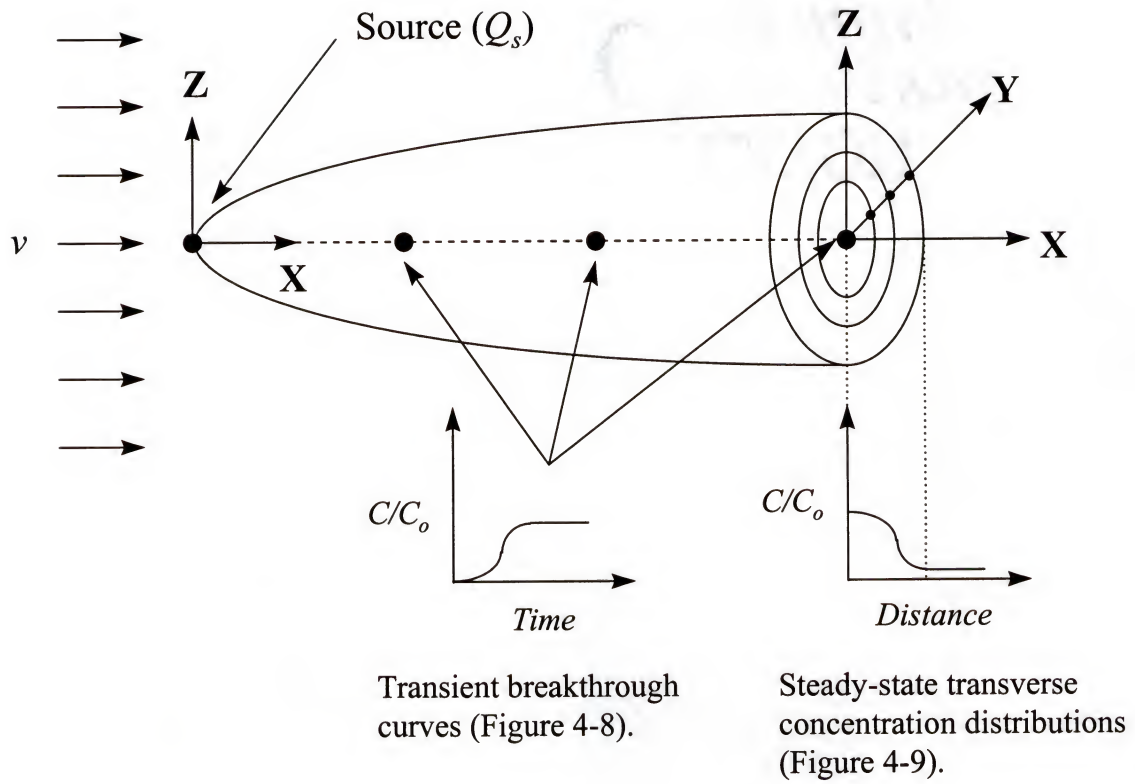


Figure 4-7. Conceptual diagram of transport characterization experiment performed to estimate longitudinal and transverse dispersivities.

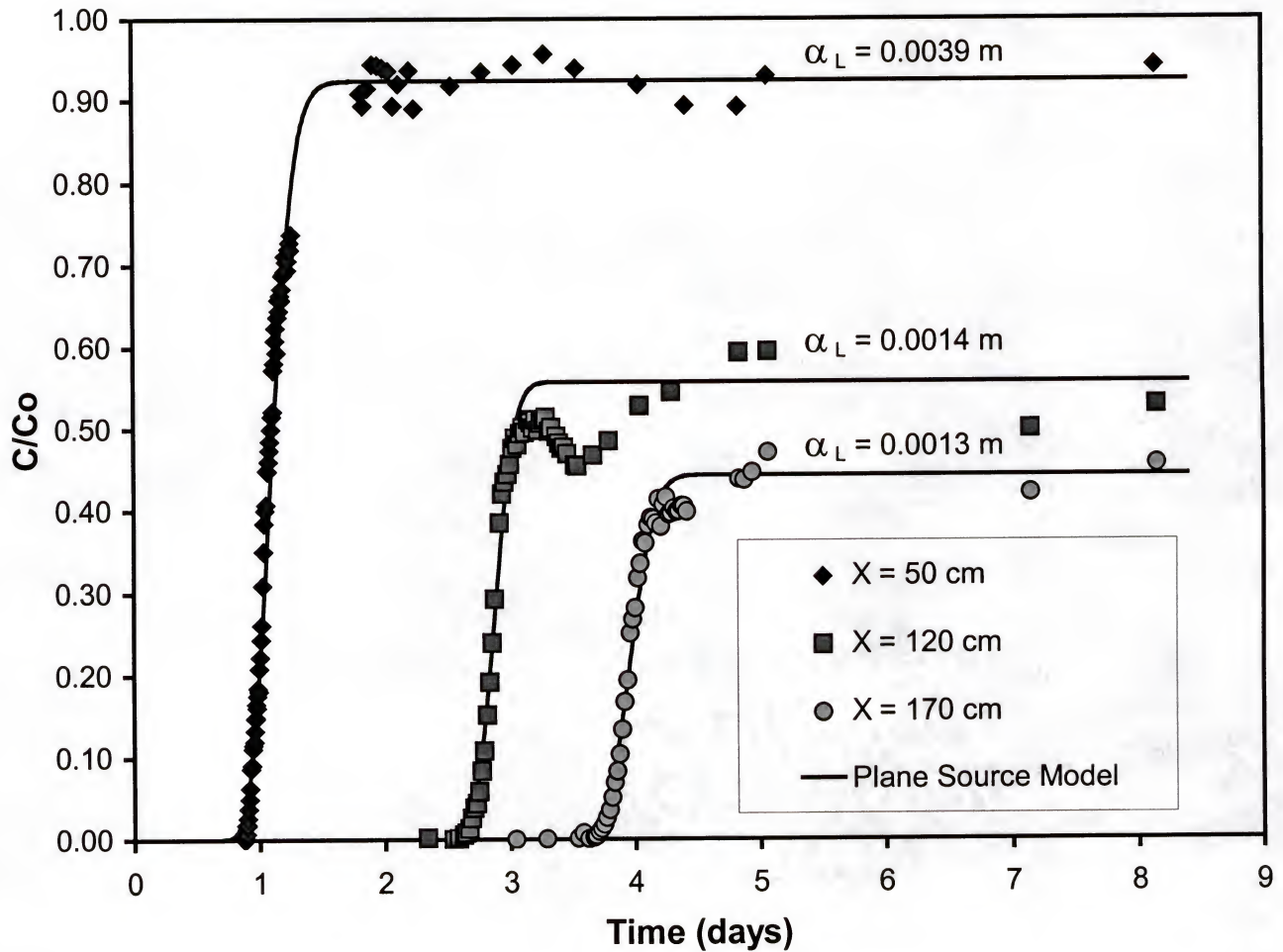


Figure 4-8. Chloride breakthrough curves ( $x = 50, 120, \text{ and } 170 \text{ cm}$ ) and the corresponding longitudinal dispersivity values determined using a Levenberg-Marquardt nonlinear parameter estimation routine.



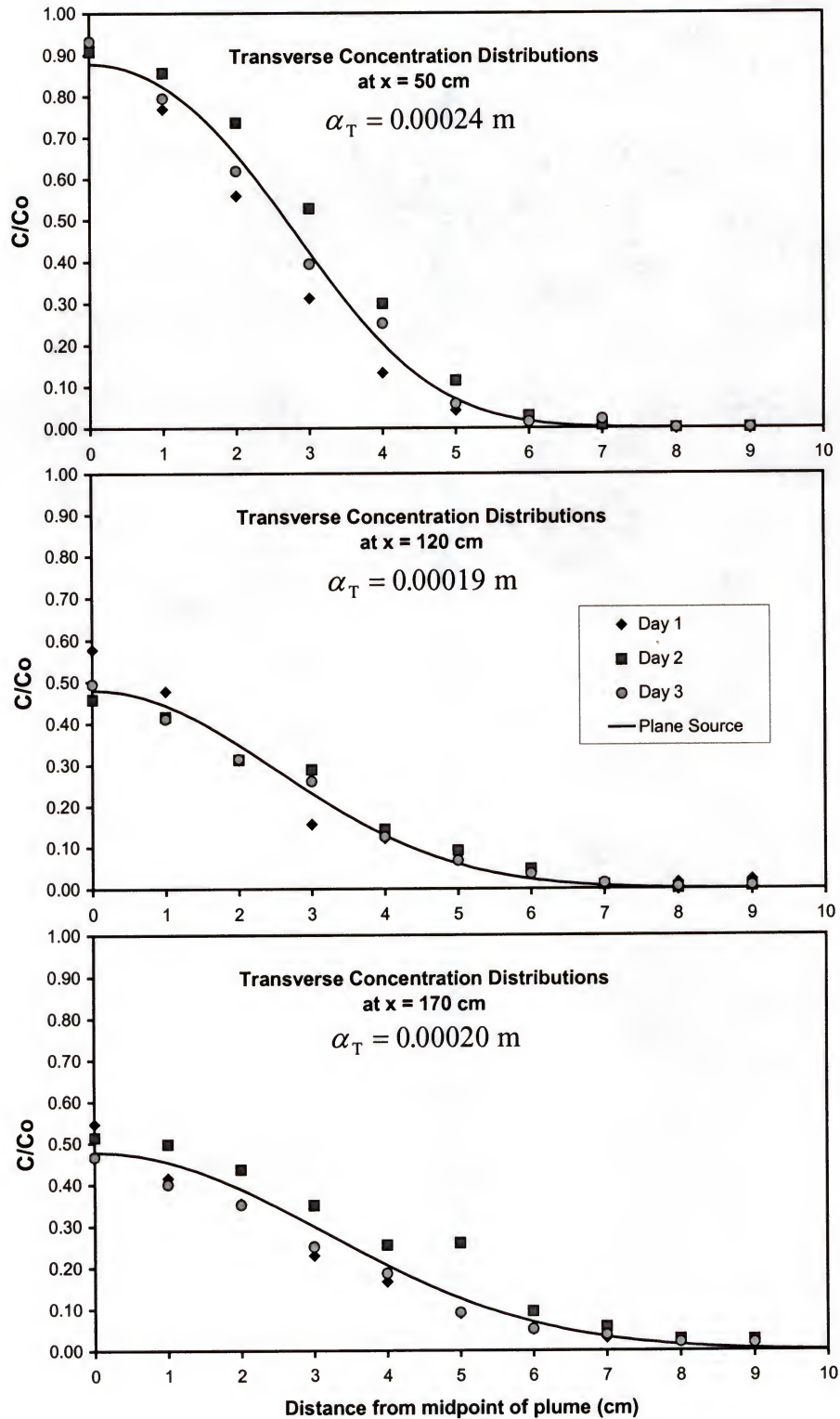


Figure 4-9. Transverse chloride concentration distributions and corresponding transverse dispersivity values determined using a Levenberg-Marquardt nonlinear parameter estimation routine.

The curves fit to the data in Figures 4-8 and 4-9 and the resulting dispersivity values were obtained using a Levenberg-Marquardt nonlinear parameter estimation algorithm (Marquardt 1963) and the analytical solution for three-dimensional transport from a continuous plane source (Domenico and Robbins 1985) as presented in equation (2-2). However, equation (2-2) was rewritten in terms of the normalized concentration  $\frac{C}{C_o}$  and by assuming that the horizontal and vertical transverse dispersivities are equal ( $\alpha_y = \alpha_z = \alpha_T$ ).

(4-6)

$$\frac{C}{C_o} = \frac{1}{8} \operatorname{erfc} \left[ \frac{x - vt}{2(\alpha_L vt)} \right] \left\{ \operatorname{erf} \left[ \frac{y + b}{2(\alpha_T x)^{1/2}} \right] - \operatorname{erf} \left[ \frac{y - b}{2(\alpha_T x)^{1/2}} \right] \right\} \left\{ \operatorname{erf} \left[ \frac{z + d}{2(\alpha_T x)^{1/2}} \right] - \operatorname{erf} \left[ \frac{z - d}{2(\alpha_T x)^{1/2}} \right] \right\}$$

where  $C$  is the plume concentration at location  $x, y, z$  and  $C_o$  is the source concentration. The dimensions  $b$  and  $d$  represent the half-width and half-height of the plane source (Figure 2-2),  $v$  is the pore water velocity,  $\alpha_L$  and  $\alpha_T$  are the longitudinal and transverse dispersivities, and  $t$  represents the elapsed time. Under steady state conditions

$$\text{As } t \rightarrow \infty; \quad \frac{x - vt}{2(\alpha_L vt)^{1/2}} \rightarrow -\infty; \quad \operatorname{erfc}(-\infty) \rightarrow 2 \quad (4-7)$$

equation (4-6) reduces to

$$\frac{C}{C_o} = \frac{1}{4} \left\{ \operatorname{erf} \left[ \frac{y + b}{2(\alpha_T x)^{1/2}} \right] - \operatorname{erf} \left[ \frac{y - b}{2(\alpha_T x)^{1/2}} \right] \right\} \left\{ \operatorname{erf} \left[ \frac{z + d}{2(\alpha_T x)^{1/2}} \right] - \operatorname{erf} \left[ \frac{z - d}{2(\alpha_T x)^{1/2}} \right] \right\} \quad (4-8)$$

In order to apply equations 4-6 and 4-8 the source dimensions  $b$  and  $d$  were estimated by equating the specific discharge of the source ( $q_s$ ) to the specific discharge of the system ( $q_x$ ). Knowing that the specific discharge is defined as the volumetric

flowrate through a given cross-sectional area, and assuming a square plane source ( $b = d$ ) the source dimension can be estimated as

$$b = \left[ \frac{Wh_s}{4} \left( \frac{Q_s}{Q_x} \right) \right]^{1/2} \quad (4-9)$$

where  $W$  is the width of the model aquifer,  $h_s$  is the average saturated thickness,  $Q_s$  is the source injection rate, and  $Q_x$  is the mean flowrate through the system.

With an estimate for the source dimensions  $b$  and  $d$ ; equation 4-8 was used with the Levenberg-Marquardt parameter estimation algorithm to determine the transverse dispersivities ( $\alpha_T$ ) corresponding to each transverse concentration distribution (Figure 4-9). The estimated values of  $\alpha_T$  were then used with equation (4-6) and the Levenberg-Marquardt algorithm to estimate the longitudinal dispersivities ( $\alpha_L$ ) corresponding to each transient breakthrough curve (Figure 4-8). The average transverse dispersivity for the three observation locations ( $x = 50$  cm,  $x = 120$  cm, and  $x = 170$  cm) was 0.0002 m, as listed in Table 4-1. It can be seen in Figure 4-8 that the longitudinal dispersivity value estimated at  $x = 50$  cm (0.0003 m) is slightly higher than the values at  $x = 120$  cm and  $x = 170$  cm (0.0014 m and 0.0013 m respectively). Inspection of the breakthrough curve at location  $x = 50$  cm reveals that the transition between the transient (rising limb) and steady state (plateau) portions of the curve was not accurately captured. It is believed that this missing information may have lead to the slightly higher longitudinal dispersivity estimate. But, rather than ignore a portion of the data set, it was decided to report the longitudinal dispersivity as a range of values (0.001 m to 0.003 m) as listed in Table 4-1.



When estimating the dispersivities, the analytical solution for transport from a continuous plane source was used rather than a solution for a continuous point source in order to provide agreement with the mathematical model presented in Chapter 2. Dispersivities estimated using a point source solution would be greater than those estimated using the plane source solution. This is because the plane source represents a larger source area than a discrete point. Tracer traveling from a point source would have to disperse a greater lateral distance in order to reach the same transverse location as tracer traveling from the extent of a plane source.

### Experiments for Validating the Mathematical Model

Once the physical, hydrodynamic, and transport characteristics of the aquifer model had been established, experiments were performed to provide a sequential set of data to test the capabilities of the mathematical model and numerical solution techniques presented in Chapters 2 and 3. The first experiment was a multiple source ionic tracer experiment and the second was a multiple source DNAPL dissolution experiment.

#### Multiple Source Ionic Tracer Experiment

The multiple source ionic tracer experiment was started by establishing a steady flow field with an average saturated thickness ( $h_s$ ) of 0.94 m, and an average pore water velocity of 0.45 m/day. Then, three steady state plumes were developed by continuous injection of bromide ( $\text{Br}^-$ ), chloride ( $\text{Cl}^-$ ), and sulfate ( $\text{SO}_4^{2-}$ ) tracer solutions (200 mg/L each) from three separate injection ports, located within the up-gradient saturated region of the flow system. Experimental conditions and parameters are summarized in Table 4-3 and the source locations are shown in Figure 4-10.

Table 4-3. System parameters for the multi-source ionic tracer experiment.

Parameter	Value
Source No. 1 (Chloride)	
Injection Point Location ( $x, y, z$ ; m)	0.20, 0.25, 0.65
Concentration (mg/L)	200
Injection Rate (mL/hour)	$1.152 \times 10^{-3}$
Source No. 2 (Bromide)	
Injection Point Location ( $x, y, z$ ; m)	0.20, 0.20, 0.50
Concentration (mg/L)	200
Injection Rate (mL/hour)	$5.76 \times 10^{-4}$
Source No. 3 (Sulfate)	
Injection Point Location ( $x, y, z$ ; m)	0.20, 0.30, 0.50
Concentration (mg/L)	200
Injection Rate (mL/hour)	$5.76 \times 10^{-4}$
Saturated Thickness at $L/2$ ( $h_s$ ; m)	0.94
Average Effluent Water Flowrate ( $Q_x$ ; m <sup>3</sup> /day)	0.0635
Specific Discharge ( $q_x$ , m/day)	0.135
Average Pore Water Velocity ( $v$ ; m/day)	0.45

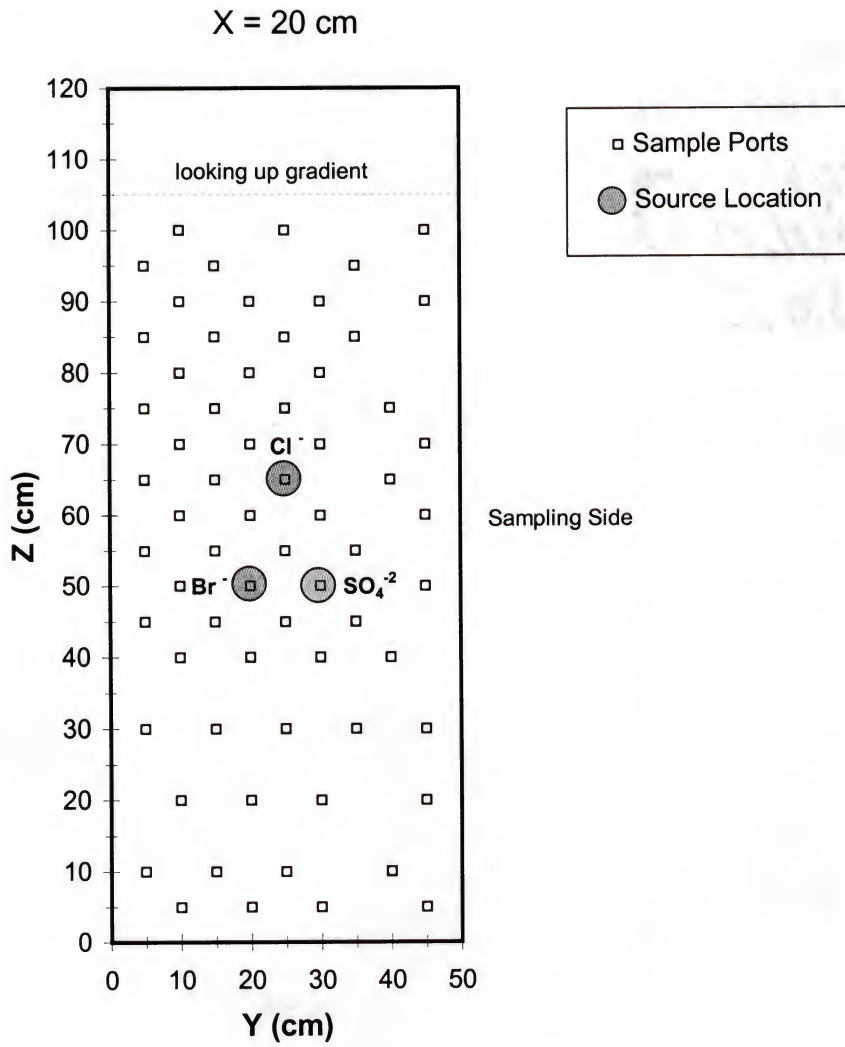


Figure 4-10. Source locations (transect  $x = 20 \text{ cm}$ ) for multiple source ionic tracer experiment.



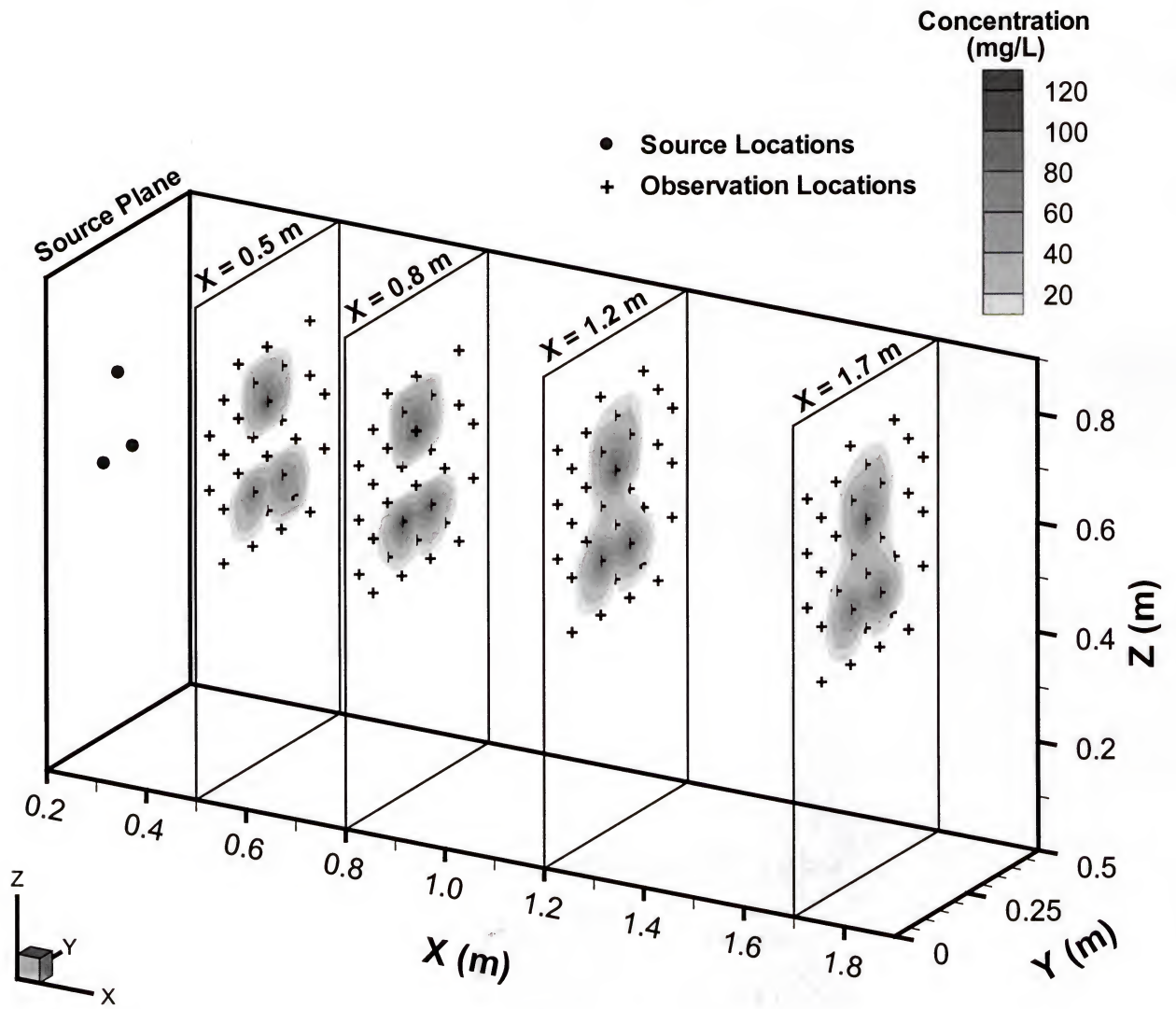


Figure 4-11. Steady state three-dimensional ionic tracer plume distribution for Day 7.

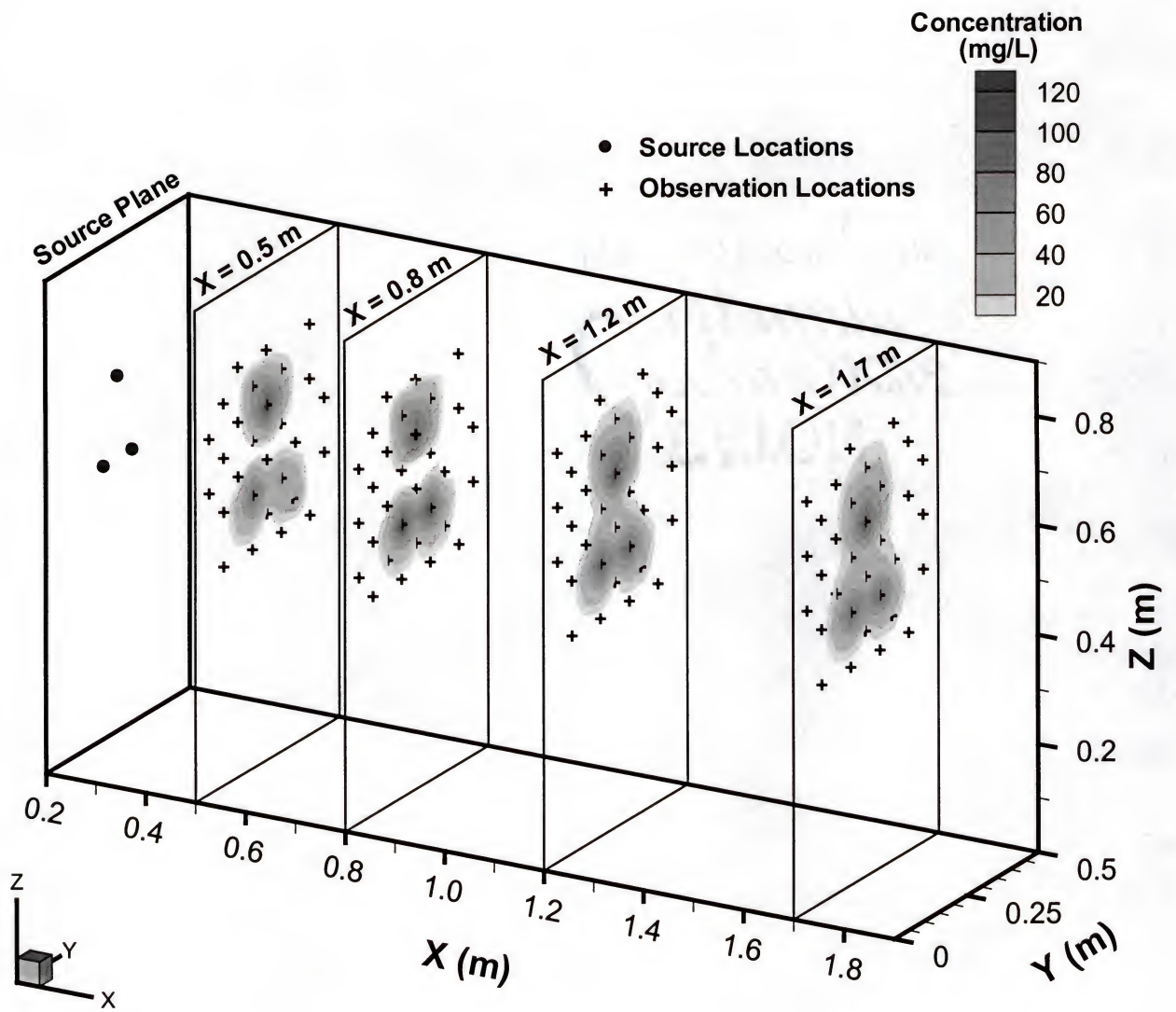


Figure 4-12. Steady state three-dimensional ionic tracer plume distribution for Day 8.

Transient tracer concentrations were measured at multiple locations along the flow path in order to determine when steady state conditions were established. Once the plumes had reached steady state (day 6), tracer concentrations were collected throughout the entire model aquifer domain on two consecutive days (day 7 and day 8). These data provided two “snap shots” of the steady state tracer distribution within the porous media and are shown in Figures 4-11 and 4-12. Inspection of Figures 4-11 and 4-12 provides a good indication that steady state conditions were established, as the plume extents and concentration magnitudes are practically identical.

The reason for using three different ionic tracers was to provide the ability to distinguish between the contributions of each point source to the combined downgradient plume. For modeling purposes the tracer concentrations can be combined and treated as one single contaminant, but having the ability to distinguish between the contributions of each source allowed for additional scrutiny of each of the numerical models capabilities to estimate the spatial distribution of mass flux. It also provided helpful information for trouble-shooting simulated flux values at zones of plume overlap during the early stages of model verification.

#### Multiple Source DNAPL Dissolution Experiment

The multiple source DNAPL experiment was started by establishing a steady flow field with an average saturated thickness ( $h_s$ ) of 0.922 m, and an average pore water velocity of 0.34 m/day. Once steady flow was established, ten DNAPL sources, composed of a mixture of 55% n-hexadecane and 45% perchloroethene (PCE) by volume, were emplaced within the up-gradient region of the flow system by slow injection of a known volume (approximately 10 ml for each source). The corresponding mole fractions were 0.3 for n-hexadecane and 0.7 for PCE. The NAPL solution also



contained the hydrophobic dye oil red-o. Experimental conditions and parameters are summarized in Table 4-4 and the source locations are shown in Figures 4-13 and 4-14.

The n-hexadecane was included in the NAPL solution for the following two reasons: to prolong the life of the PCE source zones and to help provide a visual indicator of the source zone distribution that could be recorded during excavation of the system. N-hexadecane is a hydrophobic compound, which means that it has an extremely low aqueous solubility ( $S_{n-hexa} = 3.588 \times 10^{-3}$  mg/l). By including an extremely hydrophobic compound in the NAPL mixture, a partitioning media for the PCE was created which acts to reduce the effective solubility of PCE. Reducing the effective solubility of PCE increases the lifetime of the PCE source zones. This concept can be illustrated through the use of an analog of Raoult's Law for the vapor pressure of a solution. Raoult's law states that *the vapor pressure of a solution component above a solution is equal to the product of the vapor pressure of the liquid and its mole fraction in solution* (Petrucci, 1985). The aqueous analog of this relationship can be expressed as

$$S_{i(w)}^e = X_{i(N)} S_{i(w)}^o \quad (4-10)$$

where  $X_{i(N)}$  is the mole fraction of component  $i$  in the NAPL mixture,  $S_{i(w)}^o$  is the pure aqueous solubility of NAPL component  $i$ , and  $S_{i(w)}^e$  is the resulting effective aqueous solubility of component  $i$ .

Pankow and Cherry (1996) observed that there was a wide range of aqueous solubility limits reported in the literature for chlorinated solvents such as PCE and TCE. A review of the literature found that the aqueous solubility of PCE is reported to range between 149 mg/l (Lesage and Brown, 1993) and 200 mg/L (Mabey et. al, 1982). Knowing that the NAPL source solution was composed of 45% PCE and 55% n-

hexadecane by volume, and using the range of reported PCE solubility limits, equation (4-10) can be used to estimate the expected experimental range of effective solubility for PCE. The result is that the maximum expected aqueous phase PCE concentration during the NAPL dissolution experiment should be within the range 105 to 140 mg/L. Thus, by using n-hexadecane in the NAPL solution the solubility of PCE was reduced by 70 to 50% which decreased the rate at which PCE was removed from the system and lengthened the effective life of the experimental NAPL source zones.

The second benefit of including n-hexadecane was that because it is extremely hydrophobic, it would partition onto or coat the porous media. So, even as the source zones dissolve, a portion of n-hexadecane will remain on the soil—effectively marking the extent of the initial source zones. Because the dye, oil-red-o, is also hydrophobic it will partition preferentially into the n-hexadecane, thus providing a visual indication of the NAPL source zone spatial distribution. This was done so that during excavation the source zone distribution could be recorded.

The one drawback of using n-hexadecane was that it is a light nonaqueous phase liquid (LNAPL) meaning that its density ( $\rho_{n\text{-hexa}} = 0.7733 \text{ g/ml}$ ) is less than that of water ( $\rho_w = 1.00 \text{ g/ml}$ ). In order to assure that the NAPL solution would behave as a DNAPL the solution composition was selected such that the multi-component NAPL was denser than water ( $\rho_{\text{NAPL}} = 1.16 \text{ g/ml}$ ).

Aqueous phase PCE concentrations were collected from within the flow system during the transient stage of plume development, and over an extended time-period (approximately one month) in which the plume was essentially at steady state. Once the system had reached steady state, PCE concentrations were collected throughout the entire

model aquifer domain on two occasions day 44 and day 58. Figures 4-15 and 4-16 show the measured PCE concentration distributions. Inspection of the figures indicates that the PCE plume was essentially steady state, as the plume extents and the range of concentrations are similar.

Upon completion of the experiments the porous media was excavated and the source zone distribution indicated by oil red-o dye was recorded using an excavation grid and digital photographs (Figure 4-17). The extent of the DNAPL source zones in each photograph were manually digitized and combined to develop a three-dimensional representation of the DNAPL source zones (Figure 4-18). Knowing the three-dimensional spatial distribution of the DNAPL sources will provide an accurate assessment of the mathematical models ability to determine the spatial distribution of contaminant mass flux values within a zone of multiple sources.



Table 4-4. System parameters for the multi-source DNAPL experiment.

Parameter	Value
Source No. 1	
Injection Point Location ( $x, y, z$ ; m)	0.50, 0.15, 0.75
Volume Delivered (mL)	10.78
Source No. 2	
Injection Point Location ( $x, y, z$ ; m)	0.50, 0.20, 0.70
Volume Delivered (mL)	11.78
Source No. 3	
Injection Point Location ( $x, y, z$ ; m)	0.50, 0.25, 0.65
Volume Delivered (mL)	10.45
Source No. 4	
Injection Point Location ( $x, y, z$ ; m)	0.50, 0.30, 0.60
Volume Delivered (mL)	10.44
Source No. 5	
Injection Point Location ( $x, y, z$ ; m)	0.50, 0.25, 0.55
Volume Delivered (mL)	10.15
Source No. 6	
Injection Point Location ( $x, y, z$ ; m)	0.80, 0.25, 0.65
Volume Delivered (mL)	10.44
Source No. 7	
Injection Point Location ( $x, y, z$ ; m)	0.80, 0.30, 0.60
Volume Delivered (mL)	10.44
Source No. 8	
Injection Point Location ( $x, y, z$ ; m)	0.80, 0.25, 0.55
Volume Delivered (mL)	10.44
Source No. 9	
Injection Point Location ( $x, y, z$ ; m)	0.80, 0.15, 0.45
Volume Delivered (mL)	10.44
Source No. 10	
Injection Point Location ( $x, y, z$ ; m)	0.80, 0.10, 0.40
Volume Delivered (mL)	10.44
Saturated Thickness at $L/2$ ( $h_s$ ; m)	0.922
Average Effluent Water Flowrate ( $Q_x$ ; m <sup>3</sup> /day)	0.047
Specific Discharge ( $q_x$ , m/day)	0.102
Average Pore Water Velocity ( $v$ ; m/day)	0.34

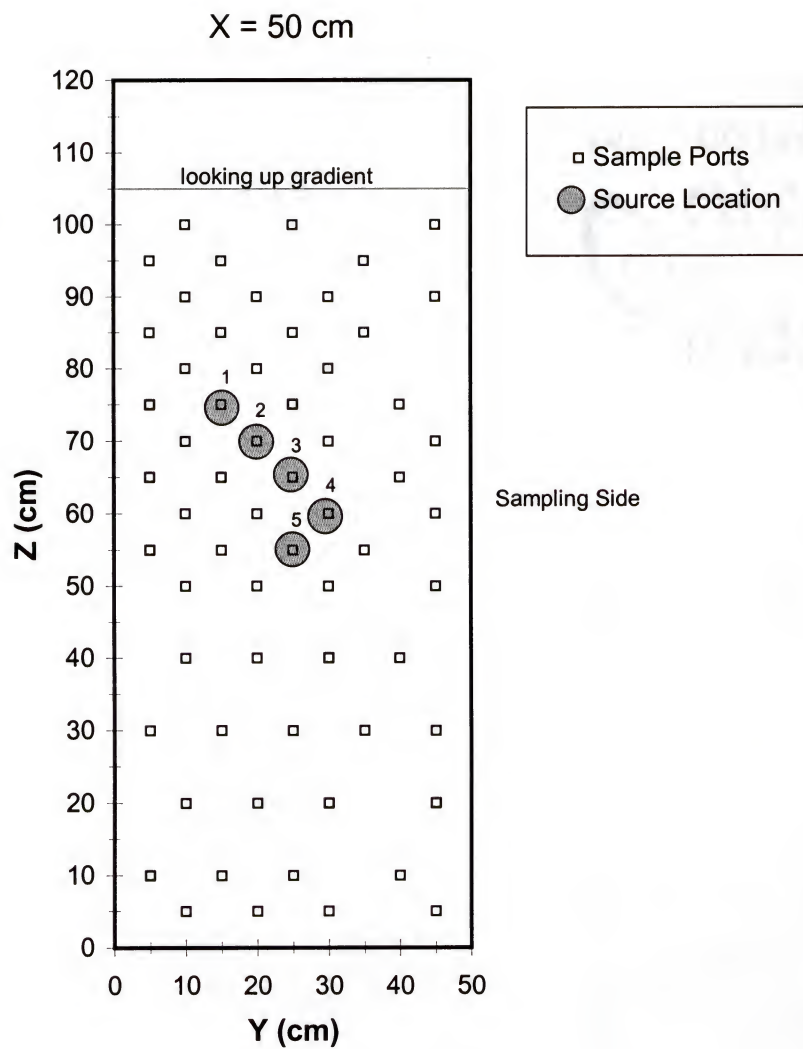


Figure 4-13. DNAPL source locations at transect  $x = 50 \text{ cm}$ .

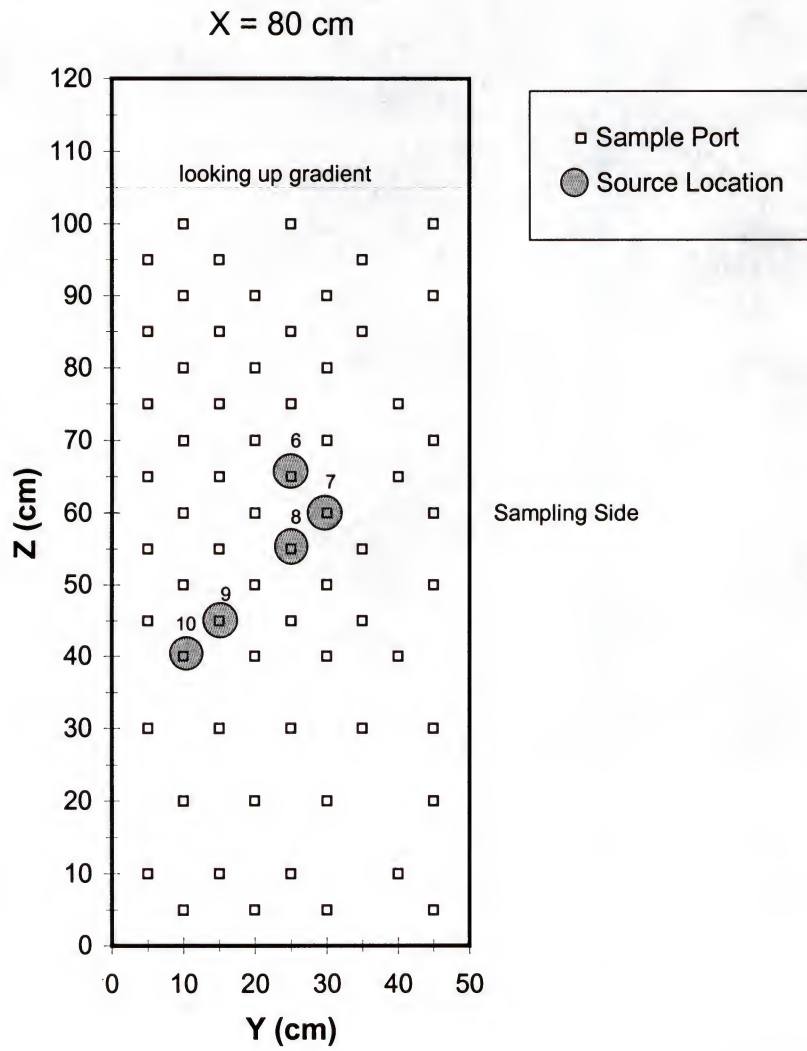


Figure 4-14. DNAPL source locations at transect  $x = 80 \text{ cm}$ .



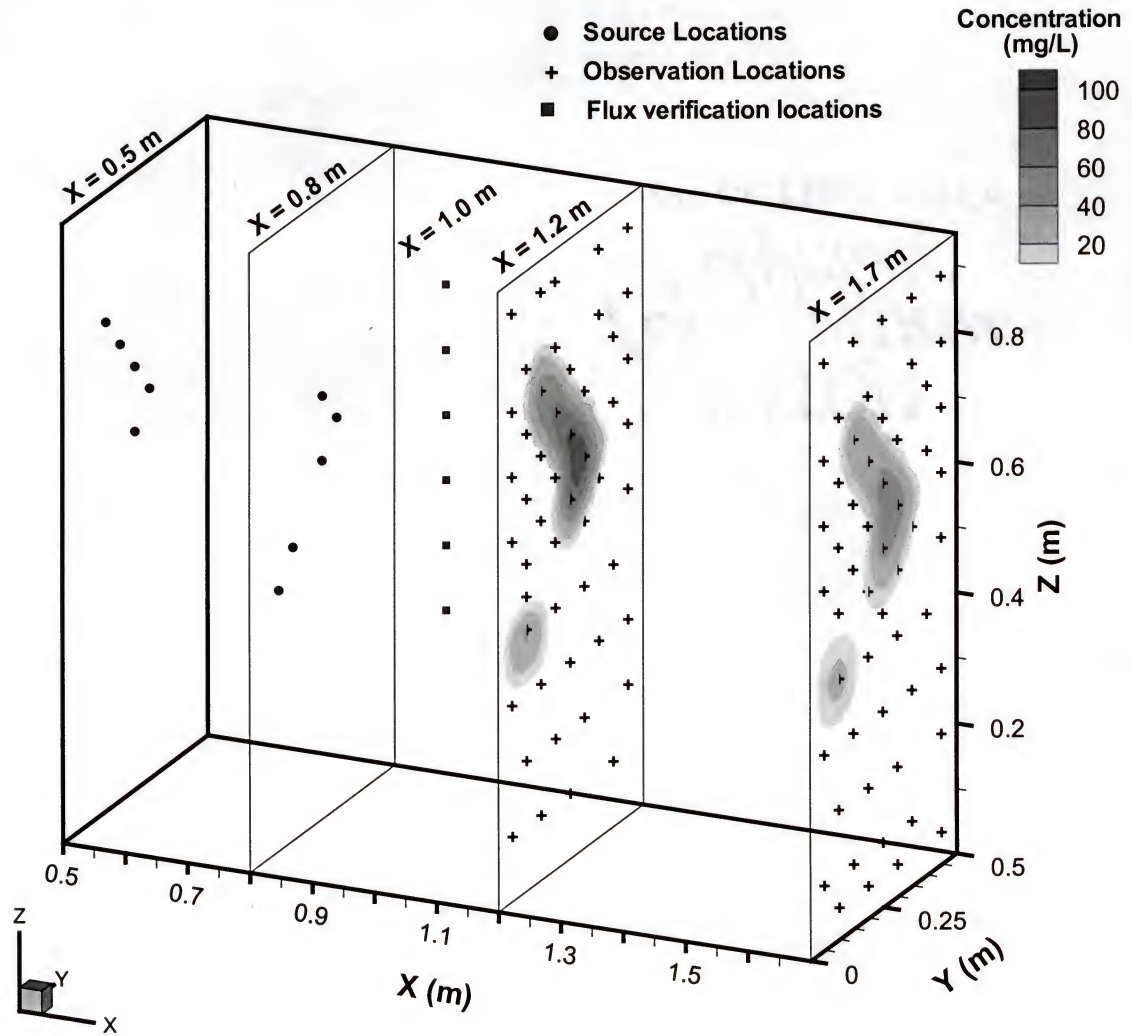


Figure 4-15. Steady state three-dimensional PCE concentration distribution for Day 44.

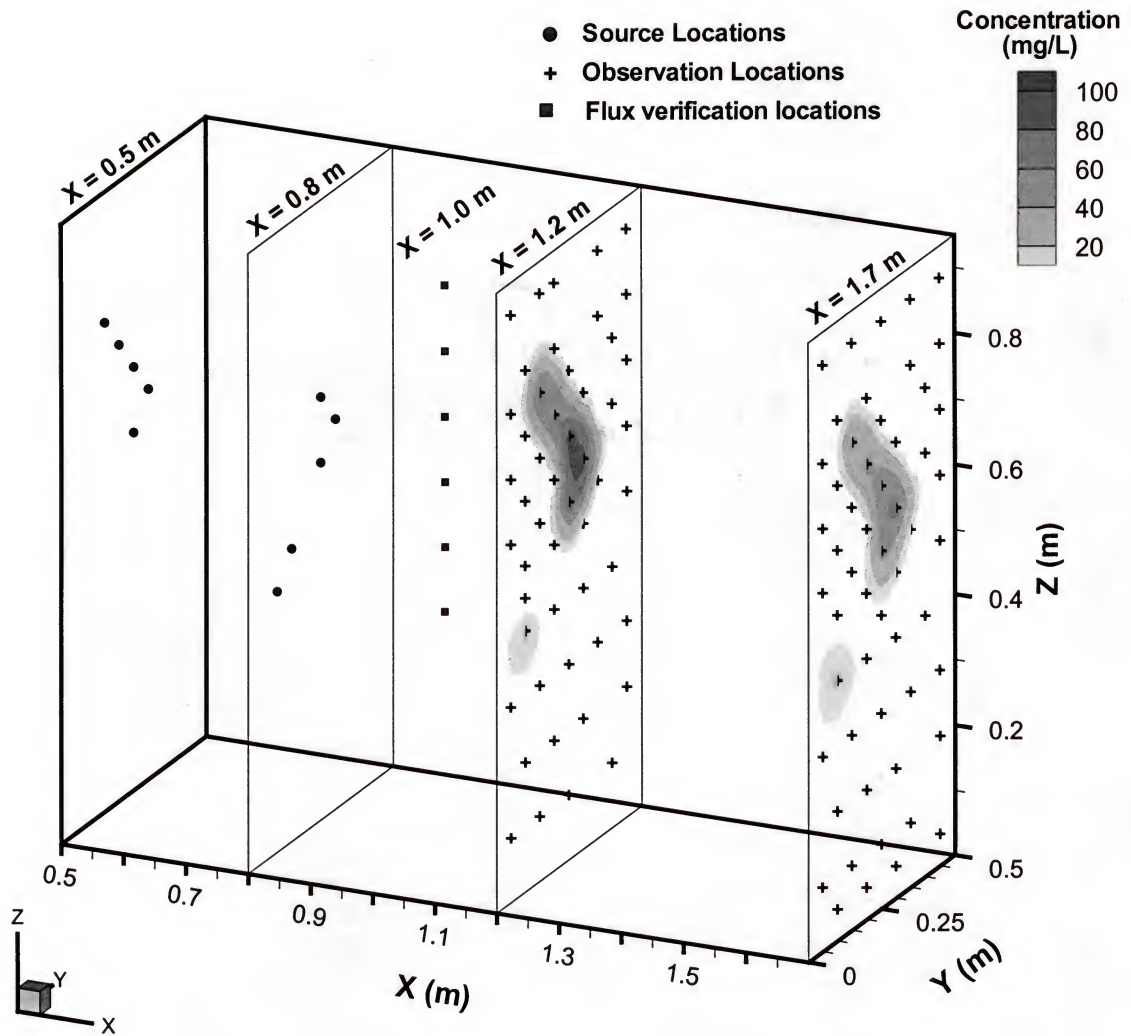


Figure 4-16. Steady state three-dimensional PCE concentration distribution for Day 58.

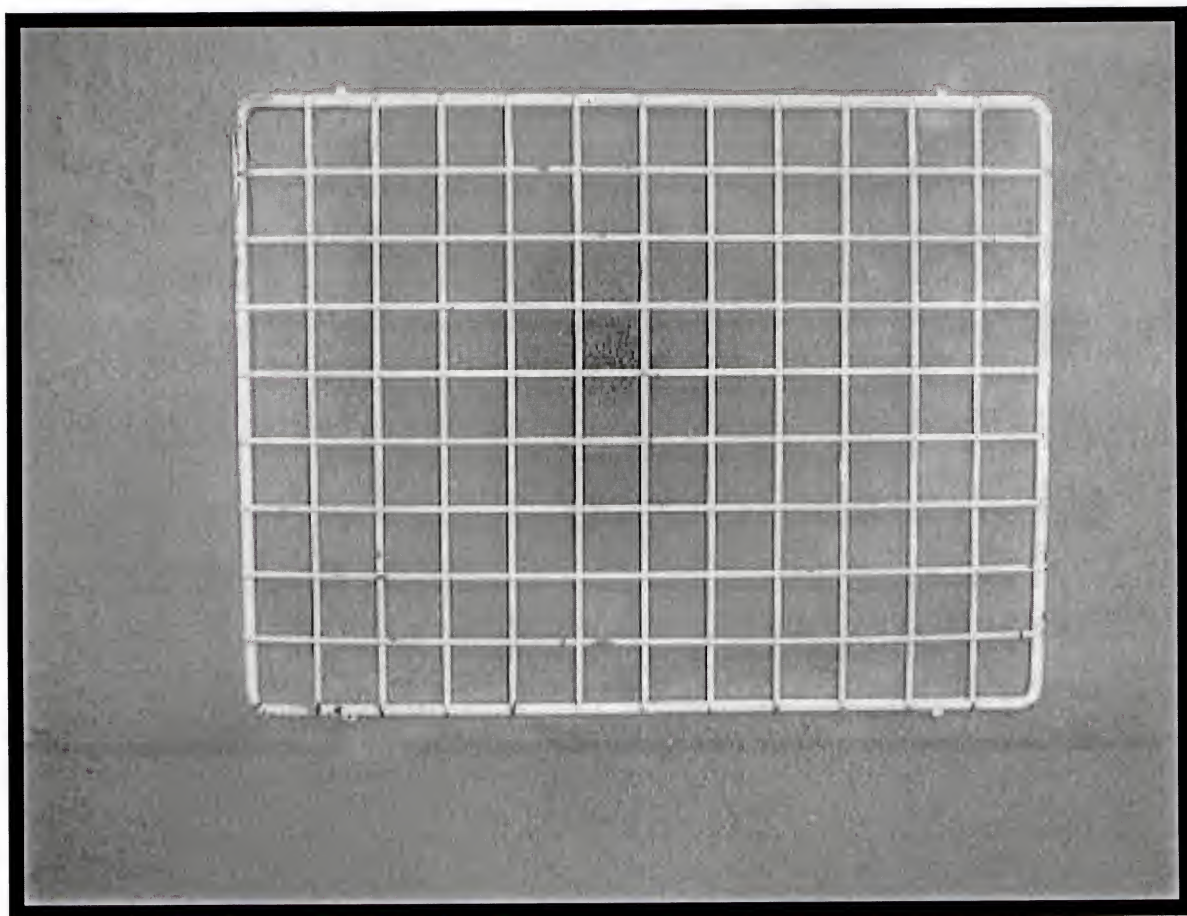


Figure 4-17. An excavation photograph showing the excavation grid (2 cm x 2 cm grid cells) and a portion of the DNAPL source zone.



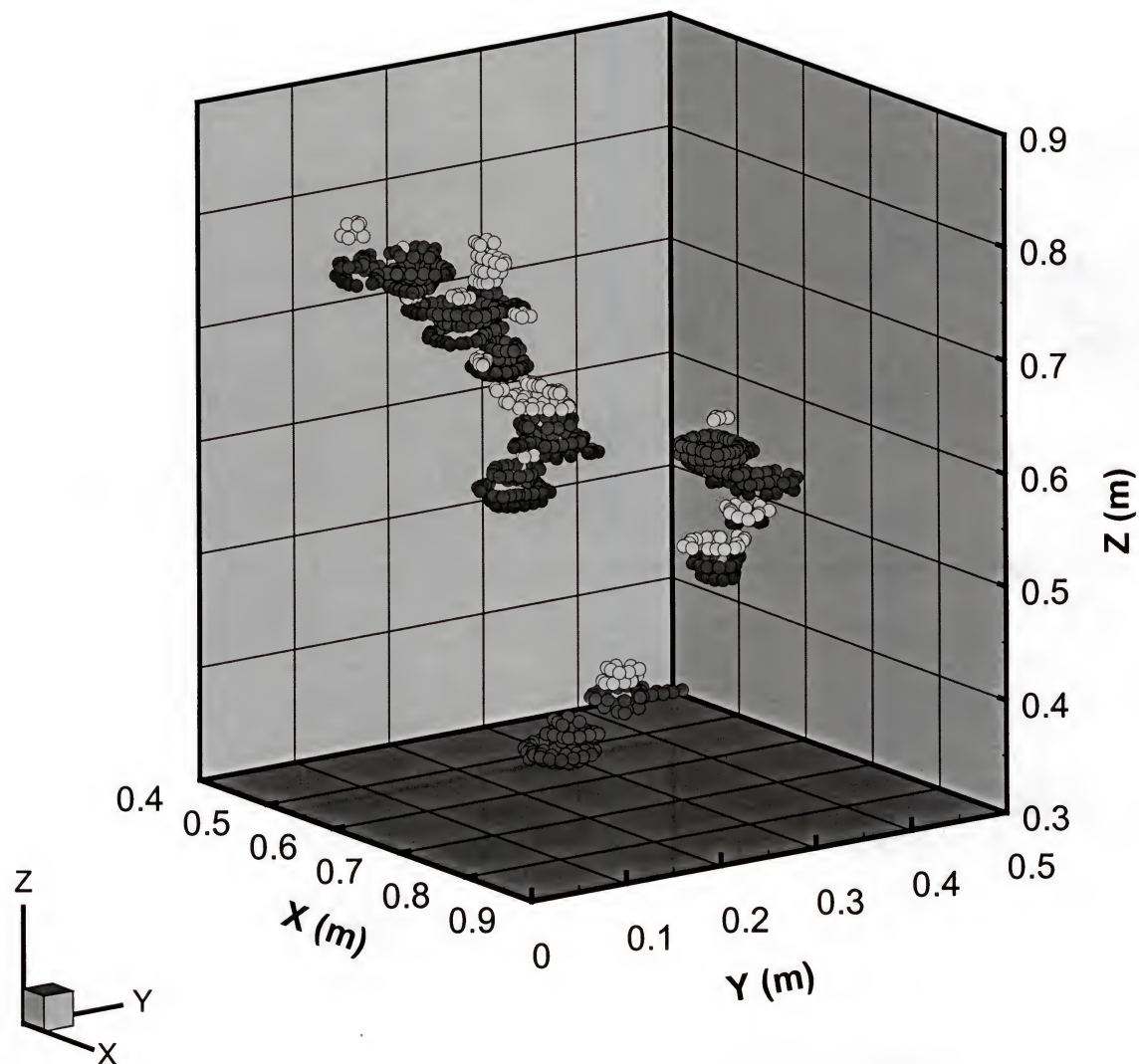


Figure 4-18. Three-dimensional representation of DNAPL source zone based upon digital excavation photographs. (The color shown matches the observed dye intensity in the excavation photographs).

## CHAPTER 5

### NUMERICAL MODELING RESULTS

The numerical modeling results are discussed in this chapter. The three solution techniques presented in Chapter 3 were applied independently in order to estimate the magnitude and spatial distribution of mass flux produced by the groundwater contaminant sources simulated in the physical experiments presented in Chapter 4.

Although each of the numerical methods applied have their own unique control parameters, there were some parameters common to all methods: the number of unknowns to be determined, number of observations, maximum number of iterations, and the convergence tolerance. In order to make an accurate assessment and comparison of each of the models capabilities, each simulation had to be performed with a consistent set of control parameters. All of the models may be capable of determining a feasible solution, but in order to make a valid comparison we want to answer questions such as which models are capable of converging on a feasible solution using the minimal amount of observation data, which model is capable of providing reliable solutions with the least stringent convergence criteria, and most importantly which model demonstrates the best computational efficiency. The process of answering these questions started by applying the individual numerical solution techniques to a simple test problem.

#### Numerically Simulated Test Problem

The test problem (Figure 5-1) was a simplified version of the intended application of the flux plane model presented in Chapter 2. A source plane consisting of 49 flux

elements was numerically simulated assuming a system specific discharge  $q = 0.5$  m/day, a transverse dispersivity  $\alpha_T = 0.0002$  m and a flux cell half-width  $b = 0.05$  m. For the test problem, three of the flux elements were set at a flux intensity of 10 mg / (cm<sup>2</sup>day) (corresponding to an initial concentration of 200 mg/l) and observed downgradient concentrations were simulated at 24 locations (12 each at distances of 0.5 m and 1.0 m). The simulated concentration observations were then used as input for the three numerical models. The goal was for each model to use the simulated observations and inversely determine the magnitude and location of the contributing flux elements. As stated earlier, the test problem is an extreme simplification, and was used primarily for algorithm debugging purposes, but it provides a good introduction to the capabilities of the numerical solution techniques.

When implementing the algorithms, each requires a set of initial guesses for the unknown variables to be determined. The goal of this study was to determine flux magnitude and spatial distribution, but because the process is based upon observed contaminant concentrations, the models were configured so that the initial guesses were supplied as concentrations and then the models used the specific discharge to determine the appropriate initial flux values.

Ideally, one would be able to input a single initial value over the entire flux plane, and this was the first approach that was attempted. The initial value for every flux cell was set as 300 mg/l (Figure 5-2A), which is equivalent to a flux intensity of 15 mg / (cm<sup>2</sup>day). Using consistent convergence criteria (1e-04) and a maximum number of iterations of 1000 simulated annealing (SA) was the only method capable of solving the problem—accurately predicting flux intensities of 10 for the appropriate cells, and zero



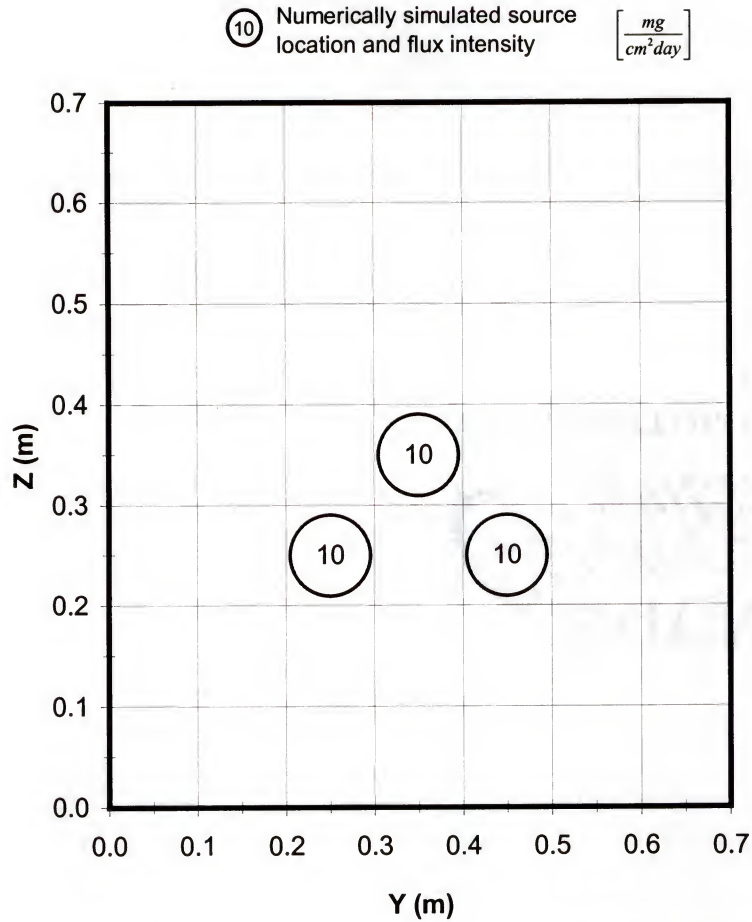


Figure 5-1. Numerically simulated flux grid used as test problem for numerical models.

System configuration:

- Specific discharge  $q = 0.5$  m/day
- Transverse dispersivity  $\alpha_T = 0.0002$  m
- Flux cell half-width  $b = 0.05$  m
- A total of 24 simulated concentration observations were provided as input (12 observations each at 0.5 m and 1.0 m downgradient of the flux plane).

○ Numerically simulated  
○ source locations

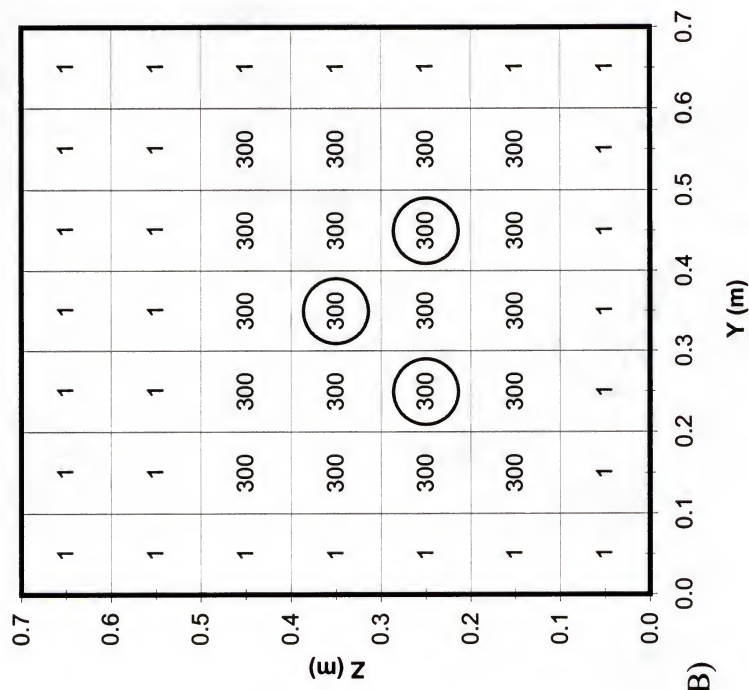
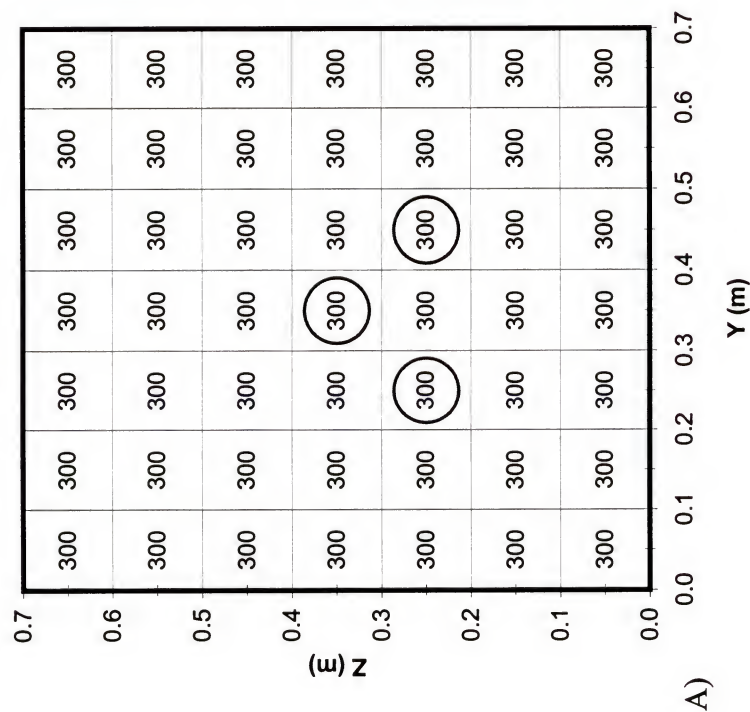


Figure 5-2. Initial Co input values for numerically simulated test problem.

(Note: With the given specific discharge of 0.5 m/day, the Co values of 300 mg/L and 1 mg/L correspond to flux intensities of 15 and 0.05  $mg / (cm^2 day)$  respectively.

everywhere else. Both shuffled complex evolution (SCE) and minimum relative entropy (MRE) reached the iteration limit and had not converged. Even with iteration limits in excess of 10,000 and more stringent convergence criteria ( $1e-12$ ) neither SCE nor MRE was capable of solving the simple system with a uniform initial input variable.

In hopes of improving the abilities of SCE and MRE to converge upon a feasible solution, the next step was to segregate the search space by providing large initial values at and around the known source cells and low initial values elsewhere (Figure 5-2B). With this configuration all of the models were capable of identifying the three source cells and accurately predicting the flux magnitude to three significant figures.

The primary reason for discussing the test problem was to demonstrate the robust search capabilities of the SA algorithm and to introduce the necessity for segregating the search space for the SCE and MRE algorithms. With the test problem, the location of each of the source cells was known and this prior information simplified the decision of how to segregate the search space. However, for application of the models to a real system, some criteria for search space segregation had to be established. This was one of the primary goals for the initial series of model applications to experimental data.

#### Multiple Source Ionic Tracer Experiment

The observed tracer concentrations from the multiple-source ionic tracer experiment performed in the three-dimensional aquifer model (Figures 4-11 and 4-12) were used as input to test the ability of each algorithm for estimating the mass flux magnitude and distribution within the system. As discussed previously in Chapter 4, the tracer concentration distribution throughout the entire system was recorded on two consecutive days (day 7 and day 8 following the start of tracer injection into a uniform



flow field with specific discharge,  $q = 0.135$  m/day). The experimental conditions are summarized in Table 4-3. The measured tracer concentrations at sampling transects  $x = 1.2$  m and  $x = 1.7$  m (Figures 4-11 and 4-12) were used as input for the flux plane models. The goal was to use these observed concentrations to inversely determine the magnitude and spatial distribution of mass flux at two intermediate flux planes ( $x = 0.8$ ,  $x = 0.5$  m) and the source plane ( $x = 0.2$  m). The experimentally measured tracer concentrations at the intermediate flux planes and the known tracer injection concentration (200 mg/l) at the source plane were used with the system specific discharge ( $q = 0.135$  m/day) to provide calculated flux values for verification of the model-simulated values.

The first step in applying the numerical models was to establish an arbitrary flux plane and discretize it into a specified number of flux elements or cells. The flux plane created for application to the ionic tracer data is shown in Figure 5-3. Each of the 442 elements represents a 2-cm square plane source with a possible source intensity ranging from a lower value of 0 mg/l to a specified upper limit. For this case, knowing that the true tracer injection concentration was 200 mg/l, a conservative upper limit was established as 300 mg/l. For application to a real system in which the source concentration is unknown, some method for estimating the appropriate upper limit will need to be established. The most conservative estimate would be determined based upon the solubility limit of the contaminants observed downgradient of the flux plane.

With the flux plane discretized and the flux cell upper limit established, the next step was to determine the initial estimates for the unknown flux intensities. As stated previously, although the goal is to estimate mass flux, the observed data were in the form

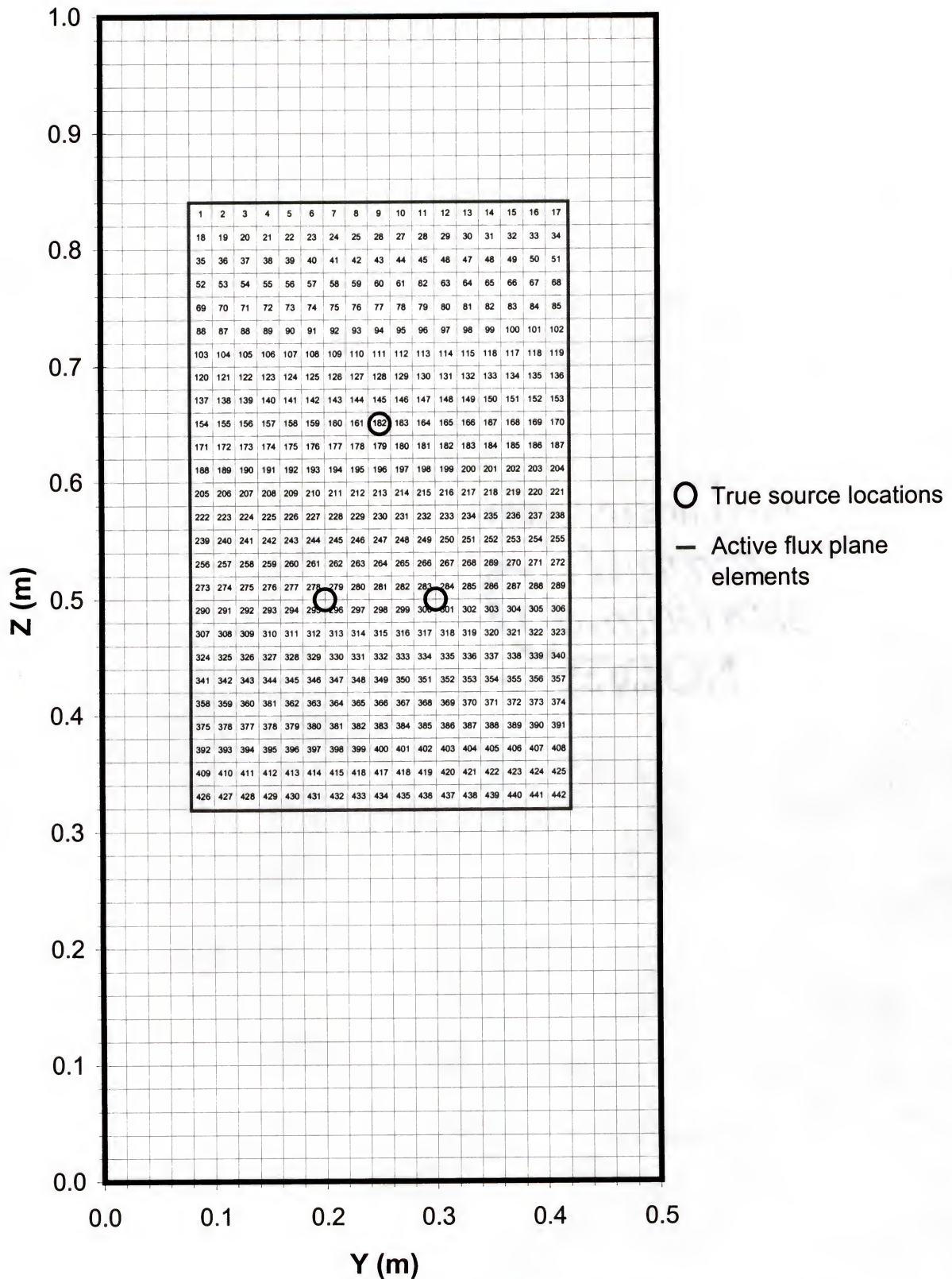


Figure 5-3. Flux plane discretization for multiple source ionic tracer experiment.

of contaminant concentrations. The models are configured so that the initial parameter values were input as expected source concentrations, which were then converted to mass flux values based upon the system specific discharge. Based upon what was learned while working with the test problem, it was known that the search space would have to be segregated in order for the numerical algorithms to converge on a feasible solution. Initially, it had been hoped that the SA algorithm would be able to solve the system using a uniform input value as it had with the test problem, but with a larger system using real data, it was not able to converge on a practical solution. It was decided that the best method for determining the initial parameter values would be to perform an order of magnitude segregation of the search space based upon the observed tracer concentrations at the furthest observation plane ( $x = 1.7$  m). Figure 5-4A shows the flux plane superimposed on the observed tracer concentration distribution at  $x = 1.7$  m. The contour plot was intentionally produced using a course contour interval so that a definite separation between high and low concentrations could be displayed. Using the 50-mg/l contour as the limiting value, the flux plane was segregated so that all the flux cells within the 50-mg/l contour were set to 50 and all those outside were set to 1. The resulting array of initial input values is shown in figure 5-4B. All of the flux cells that are shown blank in the figure were set to an initial value of 0.1 mg/l. It should be noted that the outer cells are not inactive by being set to such a low value. All of the algorithms have the ability to increase the cell values from 0.1 mg/l to the specified upper limit if necessary during the solution process. The segregation process simply helps to focus the search for a feasible solution and minimize unnecessary iterations.



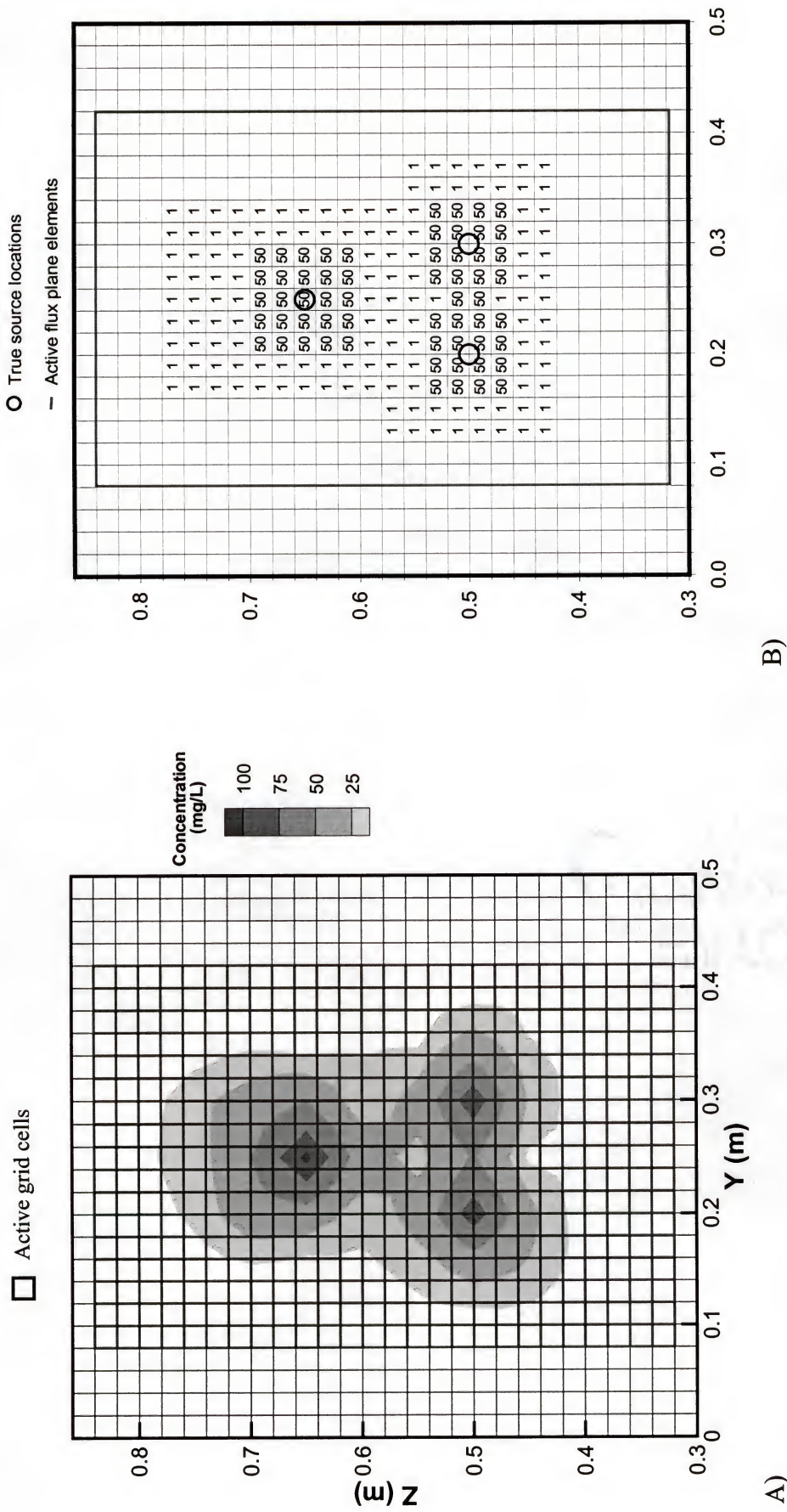


Figure 5-4. A) Flux plane grid and tracer concentrations at  $x = 1.7$  m used to segregate search space for ionic tracer experiment. B) Initial Co values used for ionic tracer experiment.

With an arbitrary flux plane established and a set of initial values provided, the models were applied using identical control parameters (convergence tolerance =  $1e-6$  and maximum number of iterations = 90,000) and 65 downgradient concentration observations (33 at  $x = 1.7$  m and 32 at  $x = 1.2$  m). First, the models were used to estimate the flux at the two intermediate flux planes ( $x = 0.8$  and  $x = 0.5$  m). The simulated flux distribution plots for the SA, SCE, and MRE algorithms are shown in Figures 5-5, 5-6, and 5-7 respectively. There are some minor differences between the simulated flux distributions at the outer extents of the plumes, for instance SCE and MRE predict slightly higher flux values in the region between the lower two sources than SA. But, in the regions of greatest flux all of the simulated results are very similar. Comparison of the simulated flux distributions and the observed concentration contour plots (Figure 5-8) indicates that each of the models were capable of matching the simulated flux distribution to the observed concentration distributions. But, this is to be expected based upon the search area segregation performed while assigning the initial parameter values. It should be noted however that the segregation does not limit the search; segregation simply provides a starting point. A good indication of this is apparent in Figure 5-5A, where the SA algorithm increased the value of three of the flux cells in the vicinity of the  $y = 0.2$  m,  $z = 0.7$  m observation location. These flux cells were within the lower value region of the initial order of magnitude segregation (Figure 5-4B) but the model was able to raise the values as needed in order to match the observed downgradient concentrations. It appears that by increasing the values in those flux cells, the SA algorithm matches the outcrop in the measured concentration distribution at flux plane  $x = 0.5$  m (Figure 5-8A).



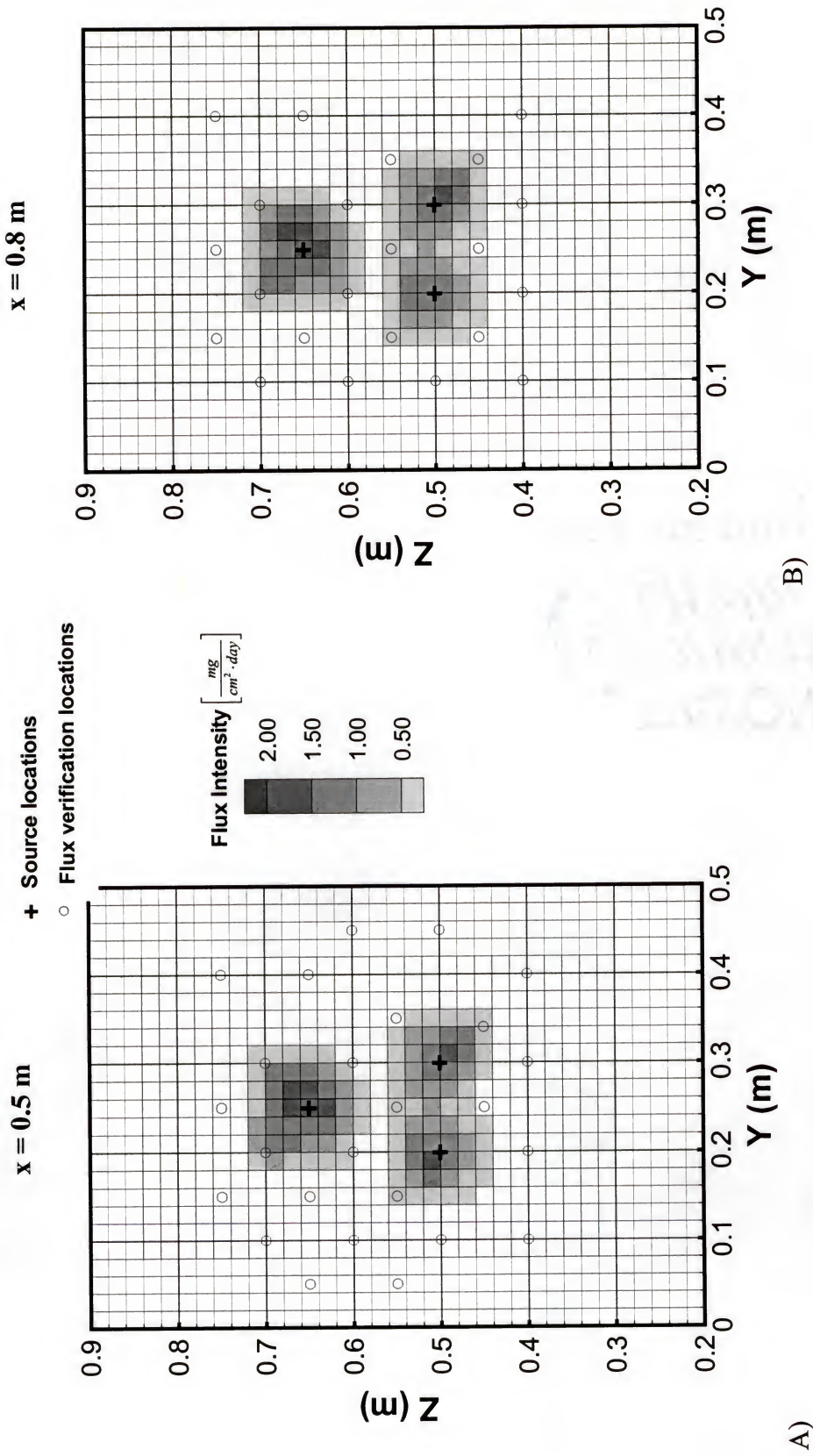


Figure 5-5: SA simulated flux distributions for Day 8.



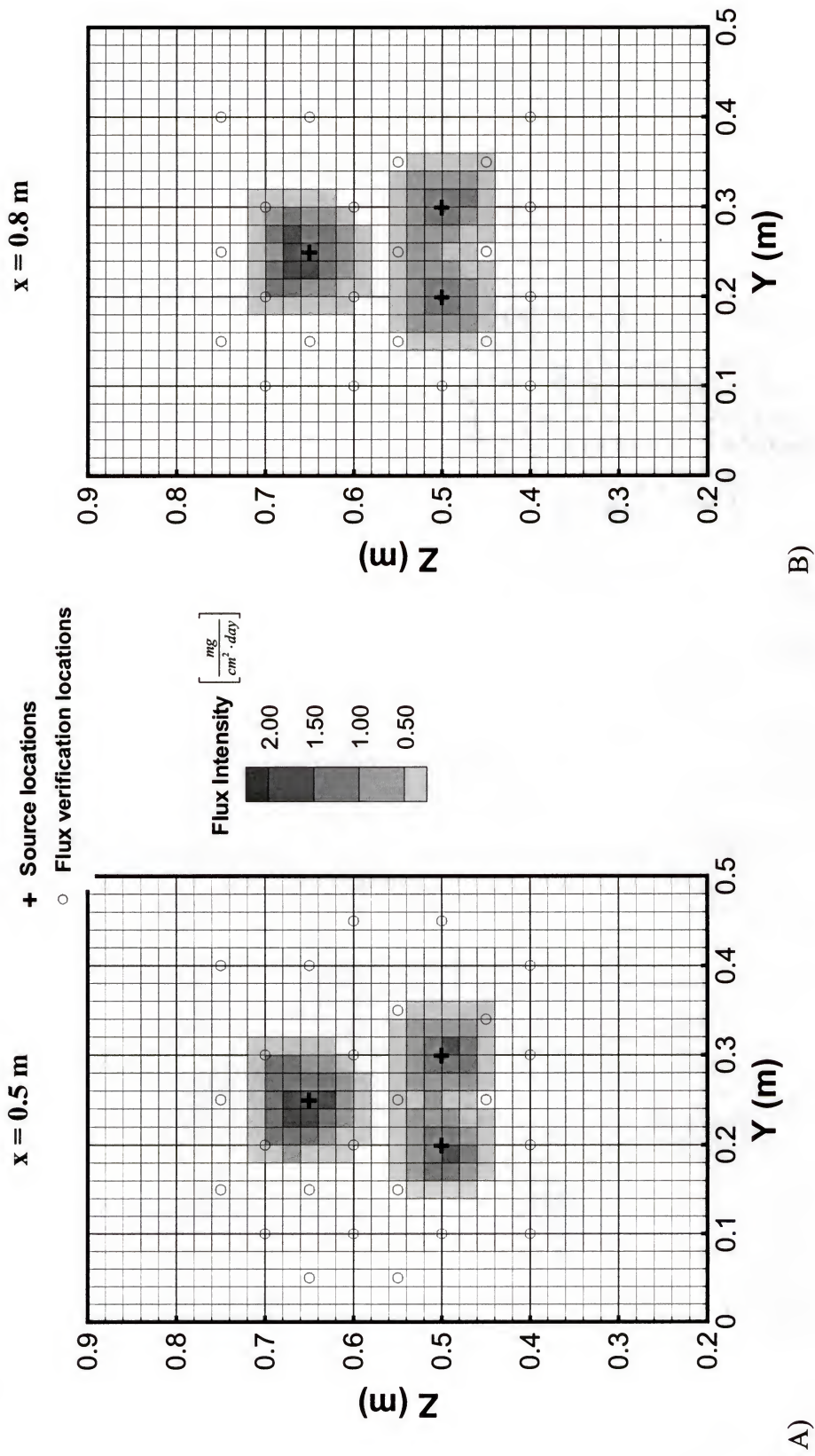


Figure 5-6: SCE simulated ionic tracer flux distributions for Day 8.

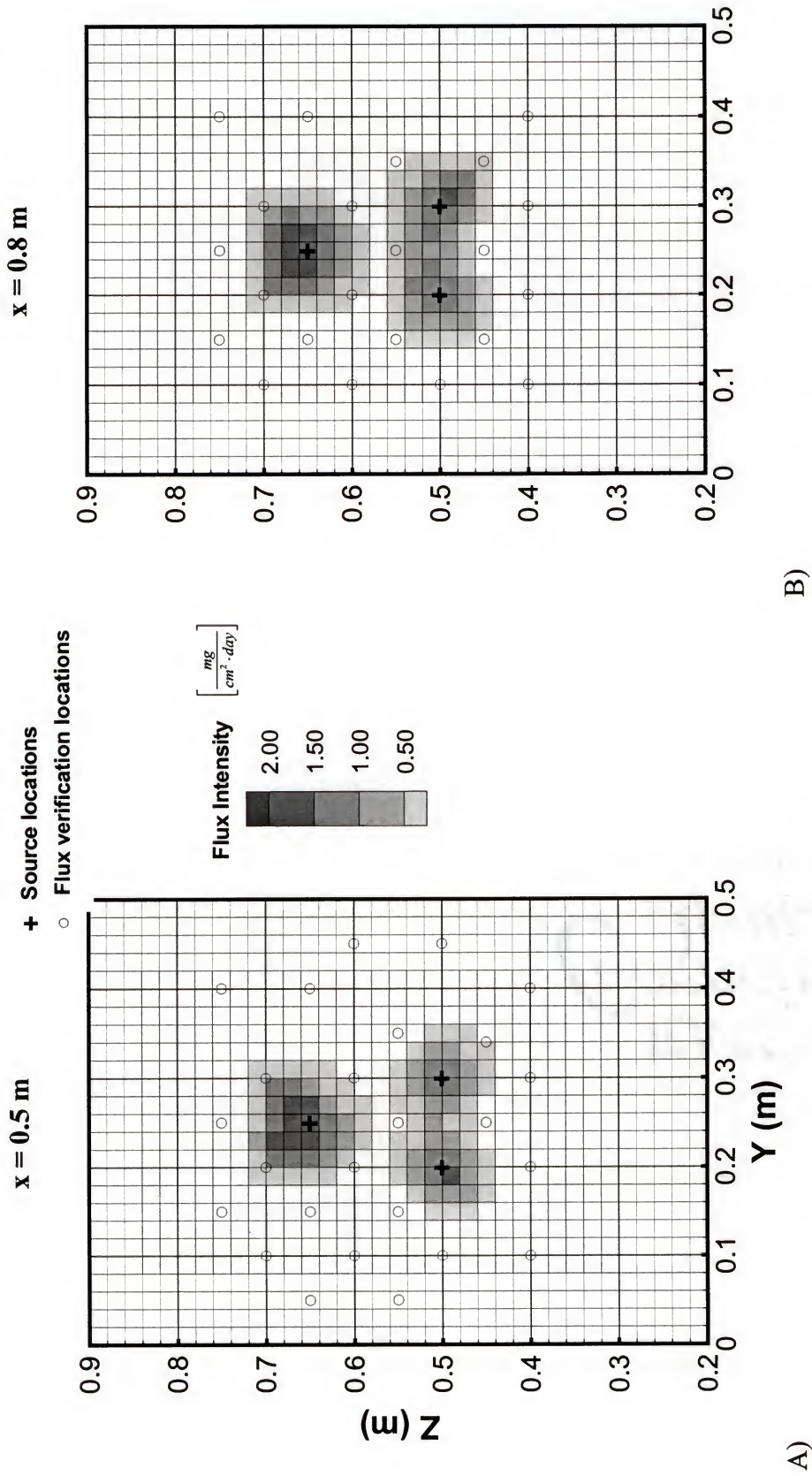


Figure 5-7: MRE simulated flux distributions for Day 8.

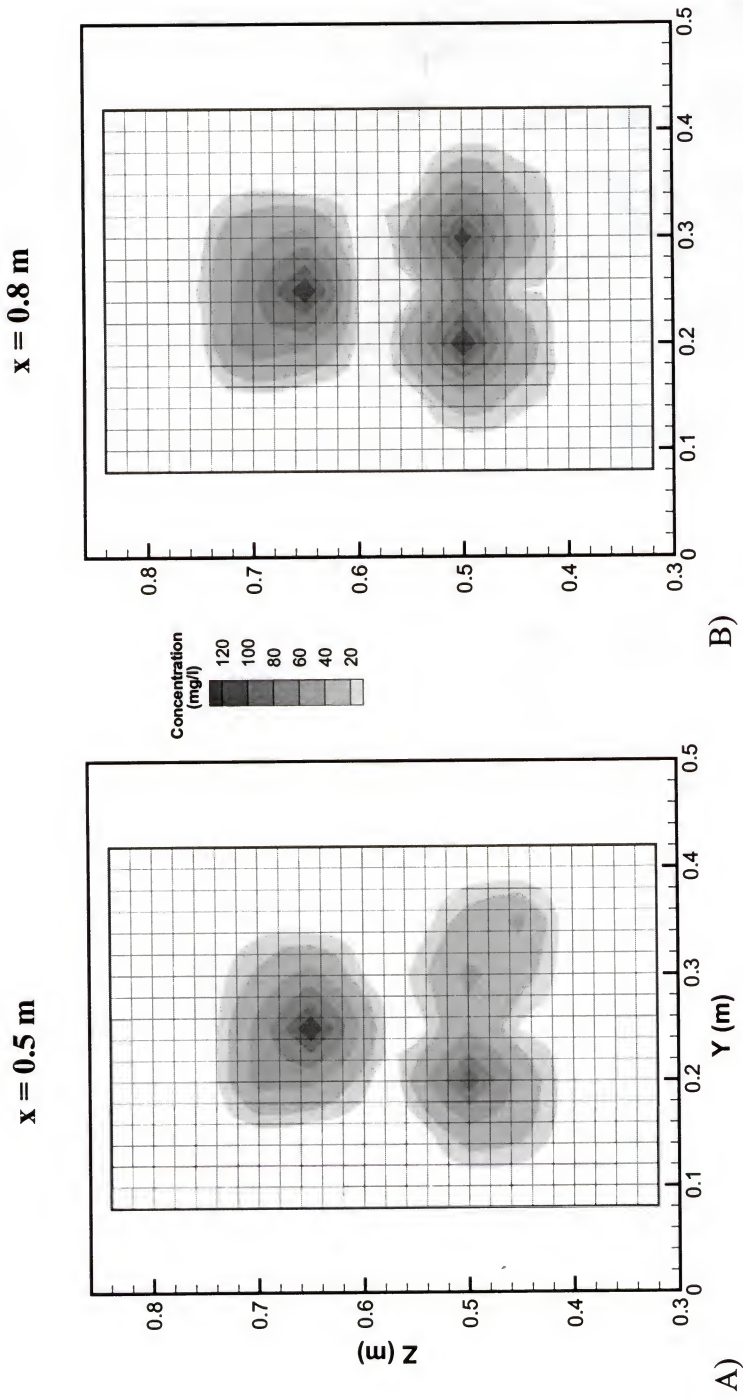


Figure 5-8: Observed tracer concentration distributions at the intermediate flux planes for day 8.



Inspection of the flux distribution plots (Figures 5-6, 5-7, and 5-8) shows that some of the flux verification locations correspond with the center of a flux plane grid cell, while others are located along the border of multiple flux cells. For the flux verifications that were coincident with the center of a cell, the model-simulated flux was taken as the flux intensity of the cell. For the flux verifications that were located on cell borders, the simulated flux was determined using linear interpolation between the flux values within each of the adjacent cells.

The model simulated and experimentally calculated flux values for Day 7 and Day 8 are shown in Tables 5-1 and 5-2 respectively. Each of the models was able to converge on a solution at both  $x = 0.8$  m and  $x = 0.5$  m for both days. However, the results were better at  $x = 0.8$  m. At locations of highest flux (corresponding to the up-gradient source locations and highlighted in Tables 5-1 and 5-2) the agreement was very good with a maximum difference of 27% for day 7 and 13% for day 8. The results are not as good at the lateral extents of the tracer plume. This is partially due to the fact that with smaller numbers (comparing 0.0 to 0.1) a small difference will still provide a large percent difference. But, part of the problem is probably a lack of observation data necessary to accurately delineate the extents of the plume.

The results were not as good for intermediate flux plane  $x = 0.5$  m. All of the algorithms had difficulty predicting the flux at location ( $x = 0.5$  m,  $y = 0.3$  m,  $z = 0.5$  m), which is directly downgradient of the bromide source. The best result was off by 184%. However, it is very likely that this is due to experimental error and not a modeling problem. The observed concentration at the location in question is peculiarly low when compared to observed concentrations directly downgradient. Also, because all of the

models consistently had problems with that location and yet did a decent job of estimating the flux at the other two source locations, it seems that experimental error would be the logical culprit. The observed values at that location are consistently low for both days, and it is likely that there was a problem within the physical system such as a partially blocked sample port.

Including the large discrepancy at flux plane  $x = 0.5$  m, the results of the three models are very similar, with the mean absolute error for each model being practically the same. It should be noted that even though there are significant discrepancies between the experimentally calculated and model-simulated values at some locations, the mean absolute error values are very small due to the numerous flux values that are less than 1 (at the extents of the plumes). Regardless of the magnitude of the mean absolute error, as long as they are determined in a consistent manner they still represent an accurate comparison of the modeling results.

When comparing the independent algorithms, a key point to note is the considerable difference in computational run time (Tables 5-1 and 5-2). SA is by far the most computationally efficient algorithm. With an average runtime of 36 seconds (0.6 minutes) SA converged 333 times faster than SCE and 52 times faster than MRE, while providing essentially the same flux estimates. Based upon computational efficiency, SA would seem to be the superior algorithm by far. However, SA does not have the capability to provide an indication of the uncertainty associated with the simulated flux values, like the MRE algorithm. For this reason, it was decided to combine the two techniques in order to improve the computational efficiency of the MRE algorithm, while still providing a measure of the reliability of the simulated flux estimates.



Table 5-1: Simulated flux values at intermediate flux planes for Day 7.

Day 7			Observed Ionic Tracer Concentration mg/L	Calculated Flux mg/(cm <sup>2</sup> day)	SA		SCE		MRE		SA-MRE	
Observation coordinates					Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference
X	Y	Z										
0.50	0.40	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.25	0.75	0.00	0.00	0.02	--	0.01	--	0.01	--	0.00	--
0.50	0.15	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.30	0.70	25.39	0.34	0.62	-0.82	0.54	-0.57	0.62	-0.81	0.60	-0.74
0.50	0.20	0.70	46.26	0.62	0.85	-0.37	0.80	-0.27	0.79	-0.26	0.80	-0.28
0.50	0.10	0.70	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.40	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.25	0.65	121.92	1.65	2.56	-0.55	2.42	-0.47	1.81	-0.10	1.77	-0.08
0.50	0.15	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.30	0.60	7.80	0.11	0.05	0.54	0.08	0.21	0.04	0.60	0.06	0.47
0.50	0.20	0.60	10.03	0.14	0.08	0.43	0.09	0.34	0.04	0.68	0.10	0.26
0.50	0.10	0.60	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.35	0.55	0.00	0.00	0.02	--	0.01	--	0.03	--	0.00	--
0.50	0.25	0.55	0.00	0.00	0.01	--	0.01	--	0.05	--	0.00	--
0.50	0.15	0.55	0.00	0.00	0.01	--	0.01	--	0.01	--	0.00	--
0.50	0.30	0.50	51.11	0.69	1.96	-1.84	1.98	-1.87	2.27	-2.29	1.98	-1.87
0.50	0.20	0.50	94.29	1.27	1.86	-0.46	1.74	-0.37	2.06	-0.62	1.82	-0.43
0.50	0.10	0.50	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.25	0.45	9.22	0.12	0.01	0.89	0.01	0.88	0.04	0.66	0.04	0.70
0.50	0.34	0.45	7.77	0.10	0.04	0.60	0.04	0.63	0.01	0.88	0.16	-0.53
0.50	0.40	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.30	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.20	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.10	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.40	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.25	0.75	9.22	0.12	0.01	0.90	0.01	0.92	0.01	0.90	0.01	0.90
0.80	0.15	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.30	0.70	32.18	0.43	0.56	-0.29	0.56	-0.30	0.59	-0.36	0.59	-0.36
0.80	0.20	0.70	59.99	0.81	0.73	0.10	0.74	0.08	0.75	0.07	0.75	0.07
0.80	0.10	0.70	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.40	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.25	0.65	119.27	1.61	1.88	-0.17	1.36	0.15	1.17	0.27	1.68	-0.04
0.80	0.15	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.30	0.60	6.14	0.08	0.06	0.26	0.09	-0.12	0.03	0.59	0.03	0.64
0.80	0.20	0.60	7.74	0.10	0.11	-0.02	0.17	-0.58	0.04	0.65	0.04	0.67
0.80	0.10	0.60	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.35	0.55	0.00	0.00	0.01	--	0.00	--	0.00	--	0.00	--
0.80	0.25	0.55	0.00	0.00	0.01	--	0.00	--	0.04	--	0.00	--
0.80	0.15	0.55	6.40	0.09	0.01	0.87	0.00	0.97	0.07	0.21	0.01	0.86
0.80	0.30	0.50	116.55	1.57	1.75	-0.11	1.69	-0.07	1.67	-0.06	1.69	-0.07
0.80	0.20	0.50	125.59	1.70	1.62	0.04	1.58	0.07	1.56	0.08	1.57	0.07
0.80	0.10	0.50	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.35	0.45	13.77	0.19	0.01	0.93	0.05	0.75	0.01	0.93	0.01	0.94
0.80	0.25	0.45	12.01	0.16	0.01	0.91	0.01	0.93	0.05	0.72	0.04	0.73
0.80	0.15	0.45	14.81	0.20	0.01	0.95	0.01	0.95	0.13	0.37	0.12	0.38
0.80	0.40	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.30	0.40	2.50	0.03	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.20	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.10	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
Mean absolute error					0.10		0.10		0.10		0.08	
Average run time (minutes)					0.62		212		32		6	



Table 5-2: Simulated flux values at intermediate flux planes for Day 8.

Day 8			Observed Ionic Tracer Concentration mg/L	Calculated Flux mg/(cm <sup>2</sup> day)	SA		SCE		MRE		SA-MRE	
Observation coordinates					Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference
X	Y	Z										
0.50	0.40	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.25	0.75	0.00	0.00	0.03	--	0.02	--	0.01	--	0.00	--
0.50	0.15	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.30	0.70	22.94	0.31	0.53	-0.70	0.51	-0.66	0.57	-0.83	0.61	-0.96
0.50	0.20	0.70	47.76	0.64	0.86	-0.33	0.73	-0.13	0.75	-0.16	0.79	-0.23
0.50	0.10	0.70	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.40	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.25	0.65	122.02	1.65	2.12	-0.29	2.03	-0.23	1.80	-0.09	2.01	-0.22
0.50	0.15	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.30	0.60	9.94	0.13	0.02	0.84	0.05	0.66	0.04	0.69	0.04	0.67
0.50	0.20	0.60	13.64	0.18	0.09	0.50	0.14	0.21	0.06	0.66	0.13	0.30
0.50	0.10	0.60	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.35	0.55	0.00	0.00	0.01	--	0.01	--	0.03	--	0.00	--
0.50	0.25	0.55	0.00	0.00	0.01	--	0.01	--	0.05	--	0.00	--
0.50	0.15	0.55	0.00	0.00	0.01	--	0.02	--	0.01	--	0.00	--
0.50	0.30	0.50	45.51	0.61	1.80	-1.93	1.74	-1.84	2.03	-2.31	2.24	-2.65
0.50	0.20	0.50	96.35	1.30	1.81	-0.39	1.68	-0.29	2.02	-0.56	2.41	-0.85
0.50	0.10	0.50	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.25	0.45	8.01	0.11	0.02	0.85	0.02	0.86	0.04	0.63	0.03	0.72
0.50	0.34	0.45	7.05	0.10	0.08	0.19	0.06	0.38	0.01	0.86	0.28	-1.94
0.50	0.40	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.30	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.50	0.20	0.40	1.07	0.01	0.00	0.98	0.00	0.99	0.00	1.00	0.00	1.00
0.50	0.10	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.40	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.25	0.75	11.95	0.16	0.01	0.91	0.01	0.92	0.01	0.93	0.00	1.00
0.80	0.15	0.75	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.30	0.70	34.69	0.47	0.51	-0.10	0.52	-0.12	0.55	-0.17	0.54	-0.15
0.80	0.20	0.70	56.73	0.77	0.70	0.08	0.72	0.06	0.71	0.07	0.70	0.09
0.80	0.10	0.70	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.40	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.25	0.65	120.46	1.63	1.42	0.13	1.59	0.02	1.76	-0.08	1.74	-0.07
0.80	0.15	0.65	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.30	0.60	5.02	0.07	0.04	0.37	0.07	-0.04	0.03	0.50	0.03	0.56
0.80	0.20	0.60	7.61	0.10	0.16	-0.53	0.19	-0.83	0.14	-0.36	0.14	-0.36
0.80	0.10	0.60	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.35	0.55	0.00	0.00	0.01	--	0.00	--	0.03	--	0.00	--
0.80	0.25	0.55	1.62	0.02	0.01	0.59	0.01	0.75	0.01	0.44	0.01	0.39
0.80	0.15	0.55	3.38	0.05	0.01	0.75	0.02	0.64	0.01	0.71	0.01	0.78
0.80	0.30	0.50	116.34	1.57	1.55	0.02	1.54	0.02	1.46	0.07	1.47	0.06
0.80	0.20	0.50	126.40	1.71	1.59	0.07	1.53	0.10	1.49	0.13	1.51	0.12
0.80	0.10	0.50	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.35	0.45	18.83	0.25	0.01	0.95	0.01	0.95	0.01	0.94	0.01	0.96
0.80	0.25	0.45	10.07	0.14	0.01	0.89	0.01	0.91	0.04	0.72	0.03	0.78
0.80	0.15	0.45	16.71	0.23	0.00	1.00	0.01	0.94	0.13	0.44	0.10	0.56
0.80	0.40	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.30	0.40	2.97	0.04	0.00	1.00	0.00	0.99	0.00	1.00	0.00	1.00
0.80	0.20	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.80	0.10	0.40	0.00	0.00	0.00	--	0.00	--	0.00	--	0.00	--
Mean absolute error					0.09		0.08		0.09		0.11	
Average run time (minutes)					0.58		188		30		6	

The result is the coupled SA-MRE algorithm. Essentially, the SA-MRE algorithm consists of passing the final simulated flux values from the SA algorithm to the MRE algorithm as initial values. This acts to improve the initial starting point for the MRE algorithm, and will usually lead to a faster solution and possibly a better result. The SA-MRE algorithm was used to estimate the flux values at the intermediate flux planes ( $x = 0.5$  m and  $x = 0.8$  m). The results are included in Tables 5-1 and 5-2 and the flux distributions are shown in Figure 5-9. The coupled algorithm did not drastically improve the simulated flux values, and for day 8 the mean absolute error was actually slightly greater than the independent SA and MRE algorithms. However, the computational efficiency of the MRE algorithm was improved from an average of 31 minutes per simulation to an average of 6 minutes per simulation. The computational efficiency was improved while still providing a feasible solution and the ability to ascertain the uncertainty associated with each flux estimate. For example, Figure 5-10 shows the SA-MRE estimates for the pdf, 95% confidence interval, and mean (expected) value of the flux intensity for flux cell 299 ( $y = .27$  m,  $z = .49$  m) at the intermediate flux plane  $x = 0.8$  m. The SA-MRE algorithm provides these estimates for each of the flux cells within the flux plane.

Finally, the three independent algorithms along with the coupled SA-MRE algorithm were used to estimate the mass flux at the tracer source plane ( $x = 0.2$  m). The known tracer injection concentration (200 mg/l) was then used to verify the flux estimates. The results are shown in Table 5-3 and the SA-MRE flux distribution plots for Day 7 and Day 8 are shown in Figure 5-11. For all of the models, the simulated flux results at the source plane were considerably better than at the intermediate flux planes



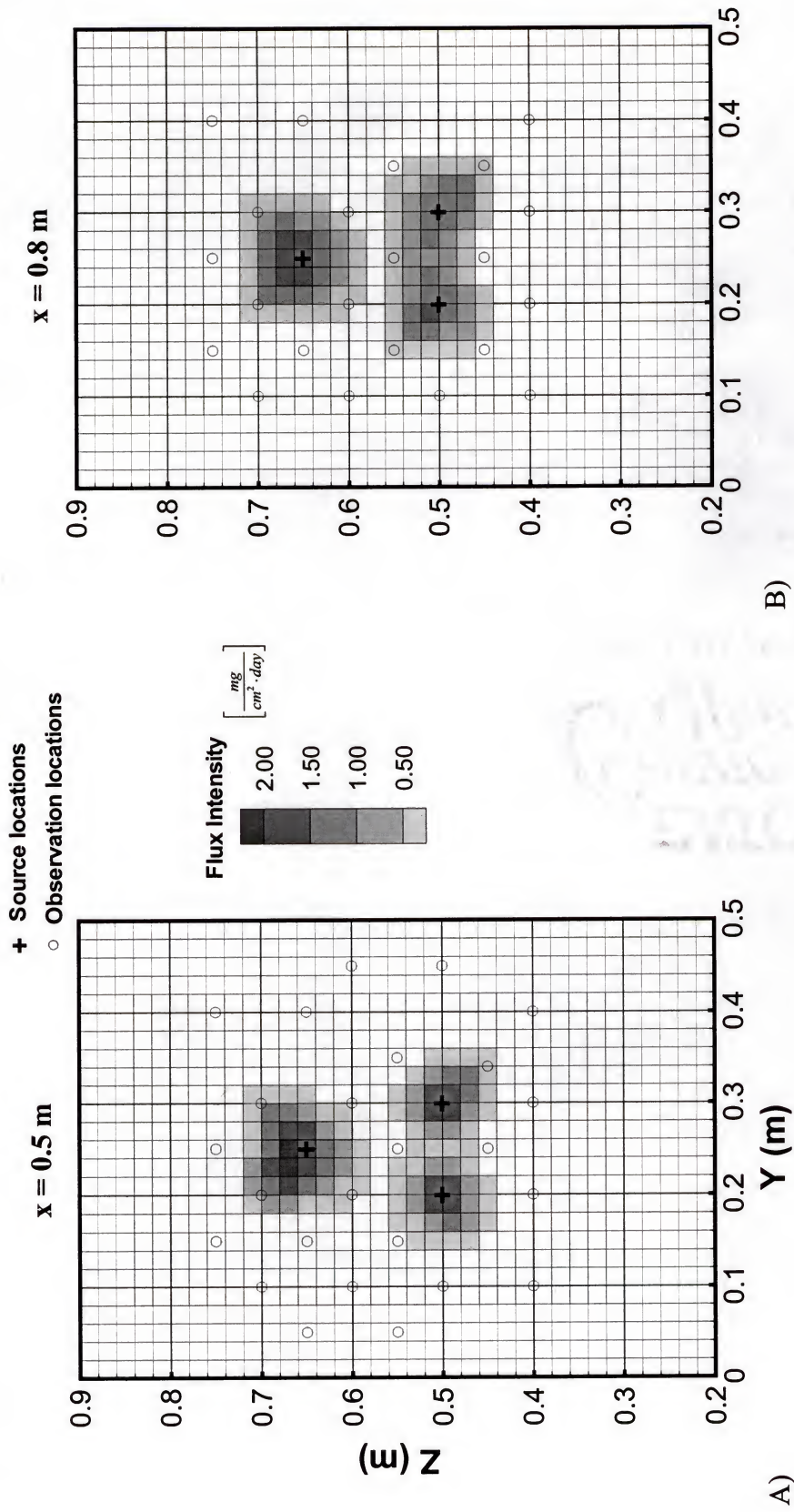


Figure 5-9: SA-MRE simulated flux distributions for Day 8



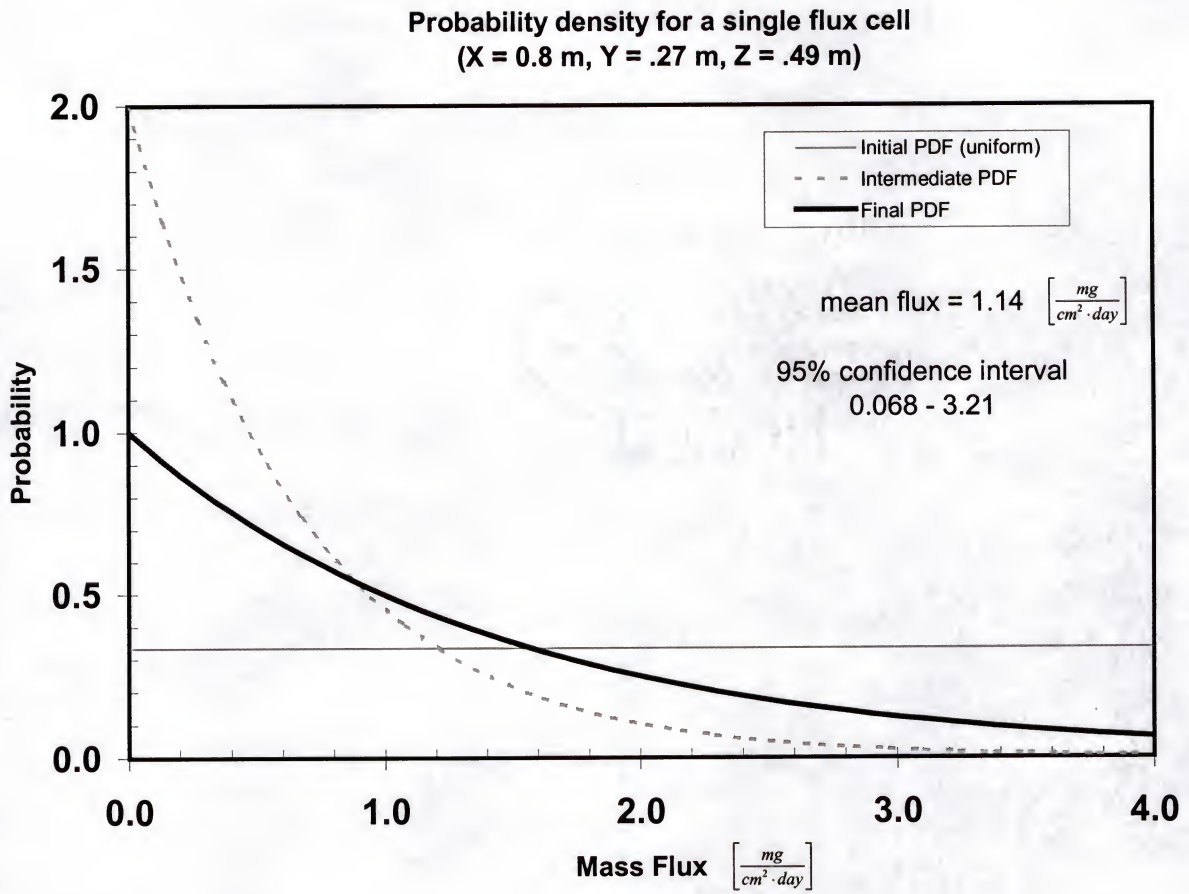


Figure 5-10. Estimated probability density function and 95% confidence interval for a single flux cell from the multiple ionic source tracer experiment.

Table 5-3: Simulated flux values at the flux plane for the multiple ionic source experiment.

Observation coordinates			Known Ionic Tracer Concentration mg/L	Calculated Flux mg/(cm <sup>2</sup> day)	SA		SCE		MRE		SA-MRE	
X	Y	Z			Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference
Day 7	0.20	0.25	0.65	200.00	2.70	3.04	-0.13	2.98	-0.10	1.98	2.45	0.09
	0.20	0.30	0.50	200.00	2.70	2.48	0.08	2.23	0.17	2.56	2.59	0.04
	0.20	0.20	0.50	200.00	2.70	2.11	0.22	1.91	0.29	2.26	2.25	0.17
Day 8	0.20	0.25	0.65	200.00	2.70	2.54	0.06	2.35	0.13	2.02	2.49	0.08
	0.20	0.30	0.50	200.00	2.70	2.01	0.26	2.17	0.20	2.25	2.28	0.16
	0.20	0.20	0.50	200.00	2.70	1.88	0.30	2.41	0.11	2.42	2.42	0.10
				Mean absolute error	0.47		0.45		0.45		0.29	
				Average run time (minutes)	0.52		198		27		5	

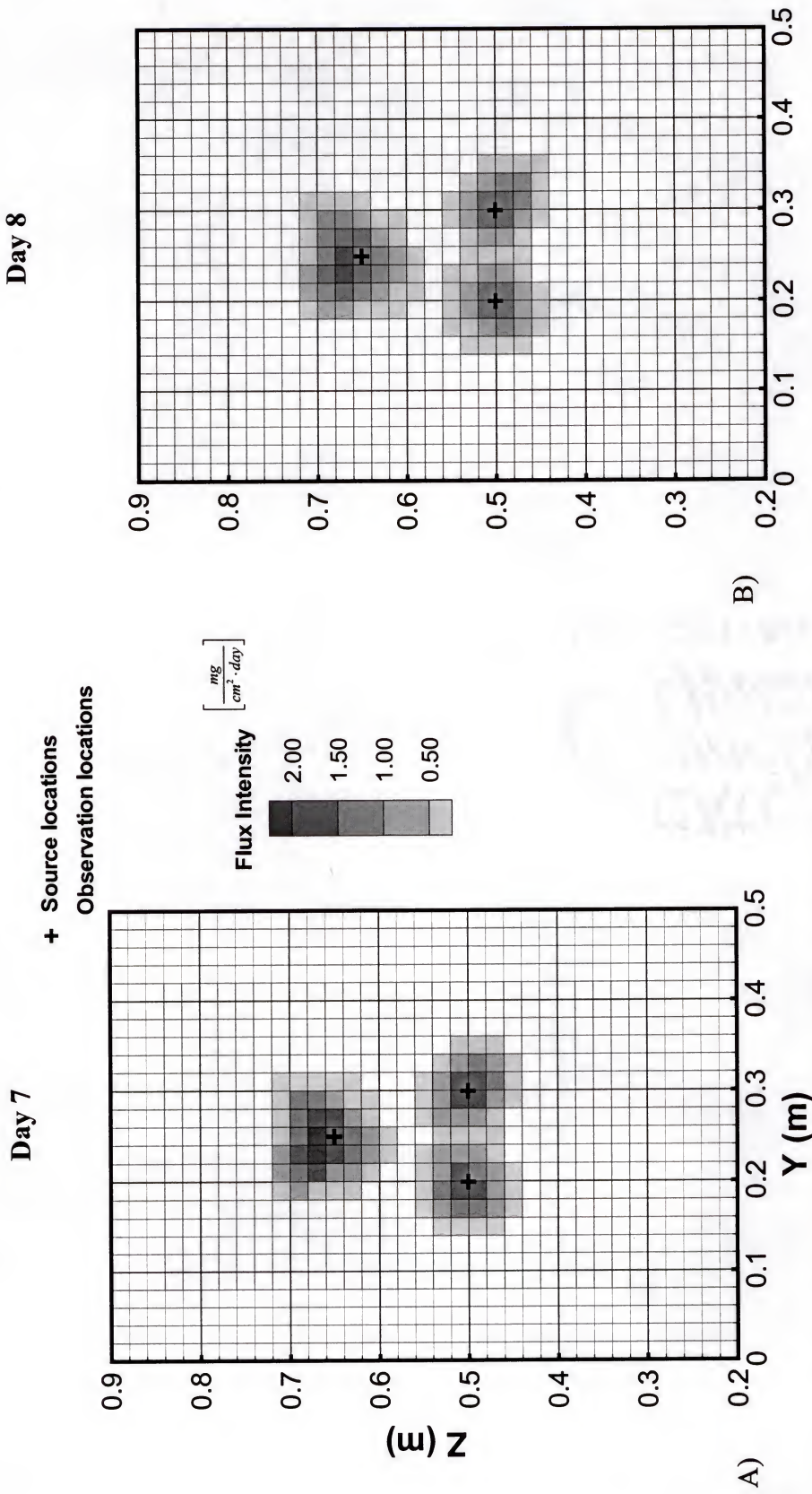


Figure 5-11: SA-MRE simulated flux distributions at the source plane for both Day 7 and 8.



with the largest deviation being 29%. The SA-MRE algorithm provided the best results improving upon both the computational efficiency of the MRE algorithm and the independent SA and MRE simulated flux values.

#### Multiple Source DNAPL Experiment

The next step in the model development process was to use results of the multiple source DNAPL experiment performed in the aquifer model (Figures 4-15 and 4-16) to further test the capabilities the flux plane numerical models. In this application the models were used to estimate the magnitude and distribution of PCE mass flux from a system of NAPL source zones.

The aqueous-phase PCE concentrations measured downgradient of the DNAPL source zones at location  $x = 1.7$  m and  $x = 1.2$  m were used as the observed contaminant concentrations. The goal was to estimate the PCE mass flux at three intermediate flux planes ( $x = 1.10$  m,  $x = 1.00$  m, and  $x = 0.90$  m) and the two sources planes (Figures 4-15 and 4-16). The simulated flux values at the intermediate flux planes were verified using experimentally calculated flux values based upon observed concentrations. The simulated flux values at the source planes were verified by comparison of the model simulated values to values determined by assuming a constant specific discharge and aqueous NAPL saturation at the flux plane. The spatial distribution of PCE mass flux at the source planes was verified by comparison to the three-dimensional source zone distribution determined during excavation of the aquifer model.

While studying the modeling results of the multiple source ionic tracer experiment, it was observed that a large amount of the flux estimation error was introduced at locations where the centroid of a flux plane element coincided with the

location of a downgradient observation point. Often times the solution algorithms would try to compensate for the observation value by altering only the coincident flux cell while neglecting the contribution of the neighboring cells. Upon closer inspection it was seen that this was not as likely to occur at locations in which observations were not coincident with the center of an up-gradient flux cell—when the observations were located on the boundary between two or more cells. Based upon these observations, it was decided to shift the flux plane for the next application, thereby assuring that none of the observation points would coincide with the center of a flux cell. Following this logic the flux plane grid shown in Figure 5-12 was established which is shifted 0.005 m from the standard axes. The grid consisted of 720 flux cells each of which was a 2-cm square plane source with a possible source intensity ranging from a lower value of 0 mg/l to a specified upper limit. For this case, the pure phase upper solubility limit for PCE, (200 mg/l) as discussed in Chapter 4, was used to determine the simulated flux upper limit (2.04 mg / cm<sup>2</sup>day ).

With the flux plane and parameter limits established the next step was to segregate the search area. Following the same procedure as in the previous section, the flux grid was superimposed on the observed PCE concentration distribution from the furthest observation plane ( $x = 1.7$  m) as shown in Figure 5-13. Using the 25-mg/l contour as the segregation contour, the initial model input values shown in Figure 5-14 were established. The flux cells that are shown as blank were given an initial value of 0.1.

Using consistent control parameters (convergence tolerance =  $1e-6$  and maximum number of iterations = 90,000) and 44 downgradient concentration observations (22 each



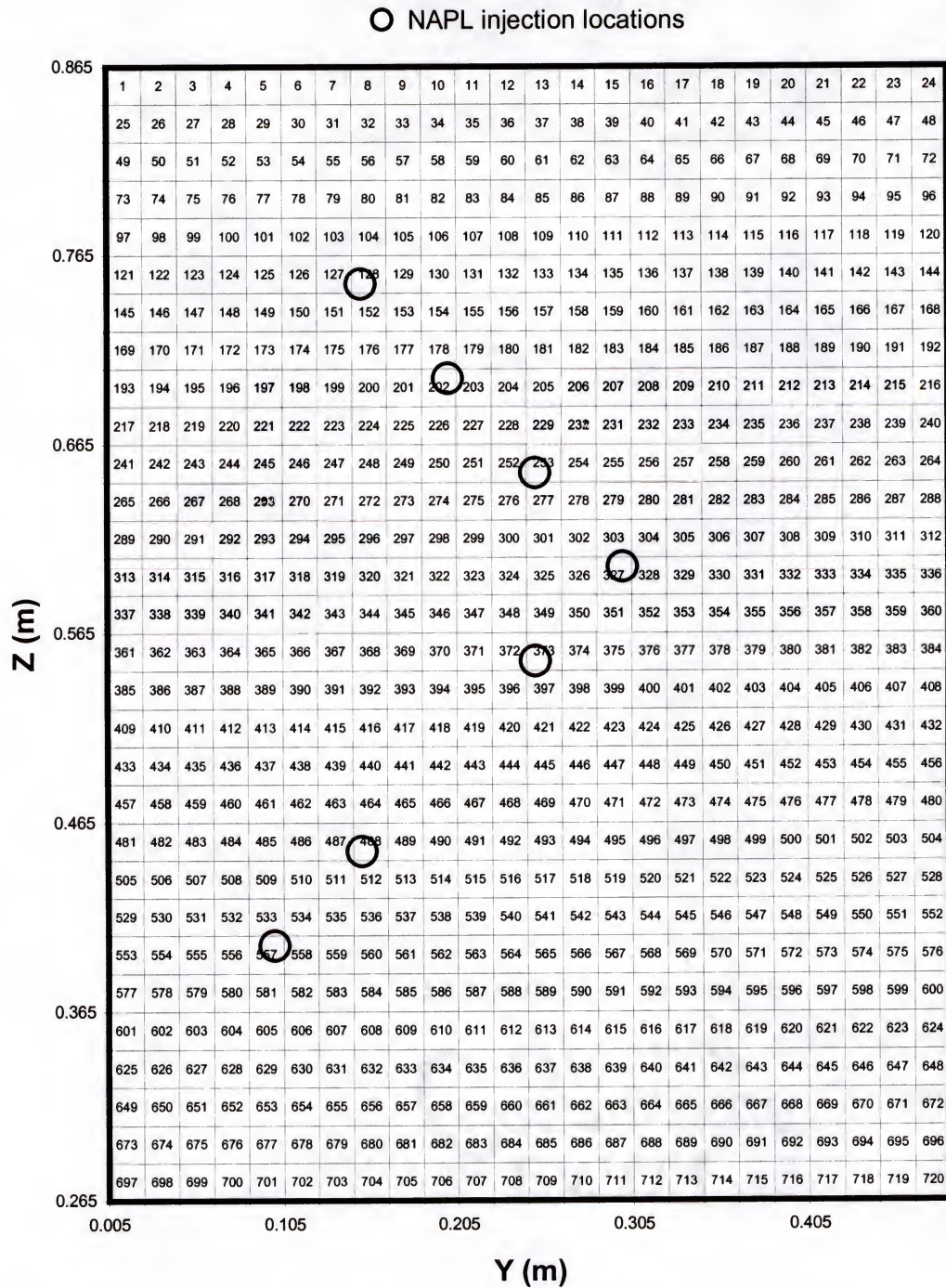


Figure 5-12. Model grid with shifted coordinates for multiple source DNAPL experiment.



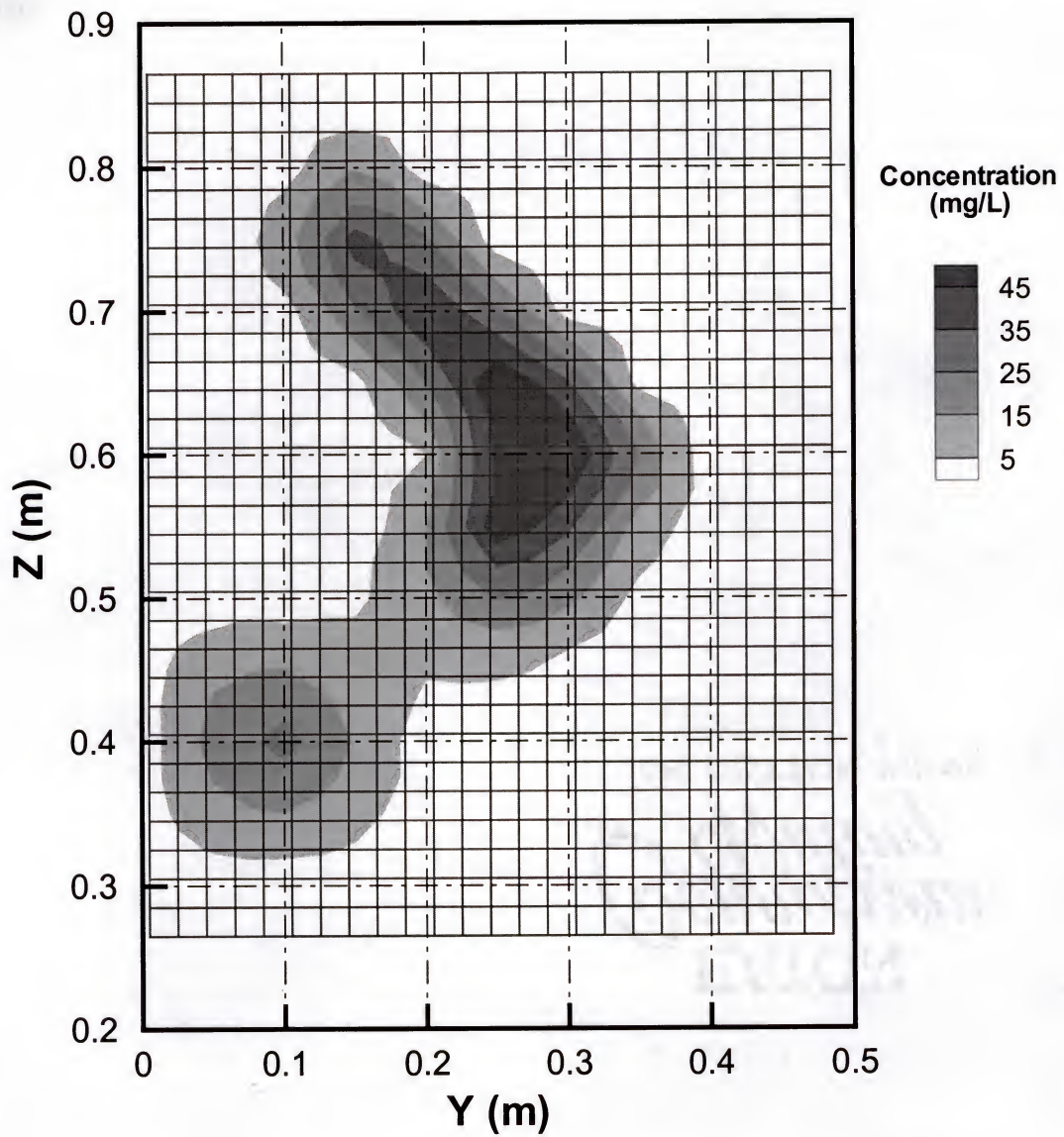


Figure 5-13. Shifted model grid with observed PCE concentrations at  $x = 1.70 \text{ m}$  (Day 44) used to establish initial model values.

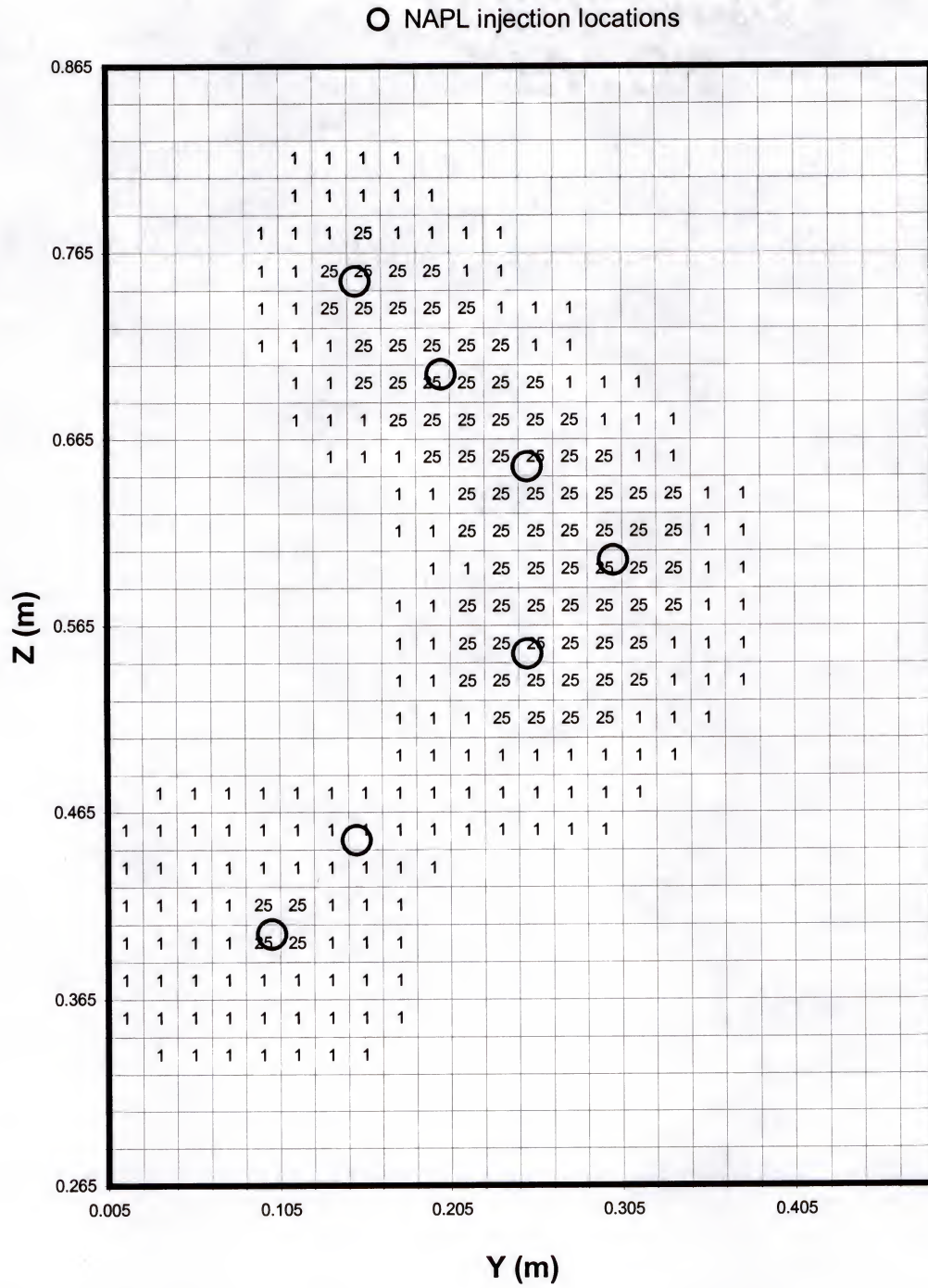


Figure 5-14. Initial Co values used for multiple source DNAPL experiment.



at observation planes  $x = 1.2$  m and  $x = 1.7$  m), the independent models along with the coupled SA-MRE model were used to estimate the PCE mass flux at the three intermediate flux planes ( $x = 0.9$  m,  $x = 1.0$  m,  $x = 1.1$  m). The model simulated flux values are tabulated and compared to the experimentally calculated values in tables 5-4 and 5-5 for day 44 and day 58 respectively. All of the models were able to converge on feasible solutions, and all the results compare favorably with the experimentally calculated values. The largest percent difference was 27% while the average absolute percent difference was 11%. Simulated fluxes using a flux plane grid that was not shifted (aligned with the standard axes) resulted in a maximum percent difference of 63% with and average absolute percent difference of 27%. This suggests that the shifted grid does help to improve the models ability to accurately simulate the flux intensities.

As with the previous application, SA is by far the most computationally efficient algorithm with an average runtime of 36 seconds (0.6 minutes), while SCE was the least efficient with an average runtime of over 9 hours. The coupled SA-MRE algorithm provided the best results for both data sets while improving the MRE computational efficiency from an average runtime of 53.5 minutes to 10 minutes while providing an estimate of the uncertainty associated with each simulated flux value. For example, Figure 5-15 shows the SA-MRE simulated mean (expected) value, 95% confidence interval, and probability density function for flux cell 373 ( $y = .255$  m,  $z = .455$  m) at the intermediate flux plane  $x = 1.1$  m.

The SA-MRE simulated flux distributions for day 44 and day 58 are shown in Figures 5-16 and 5-17. The distributions appear relatively consistent between the two days. Inspection of the plots demonstrates the expected result that as the distance from



Table 5-4: Simulated flux at the intermediate flux planes for day 44 of the multiple-source PCE dissolution experiment.

DAY 44				Observed PCE Concentration mg/L	Calculated Flux mg/(cm <sup>2</sup> day)	SA		SCE		MRE		SA-MRE	
X	Y	Z				Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference
1.10	0.25	0.85		0.39	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.10	0.25	0.75		0.38	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.10	0.25	0.65		102.04	1.04	0.91	0.12	0.87	0.16	1.11	-0.07	0.93	0.11
1.10	0.25	0.55		81.33	0.83	0.87	-0.04	0.83	0.00	1.02	-0.23	0.84	-0.01
1.10	0.25	0.45		1.52	0.02	0.01	--	0.01	--	0.01	--	0.01	--
1.10	0.25	0.35		0.45	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.00	0.25	0.85		0.37	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.00	0.25	0.75		0.30	0.00	0.00	--	0.01	--	0.01	--	0.00	--
1.00	0.25	0.65		106.07	1.08	1.03	0.05	1.06	0.02	1.10	-0.01	1.06	0.02
1.00	0.25	0.55		112.76	1.15	0.93	0.19	1.00	0.13	1.05	0.09	1.13	0.02
1.00	0.25	0.45		1.86	0.02	0.00	--	0.01	--	0.00	--	0.01	--
1.00	0.25	0.35		0.51	0.01	0.00	--	0.00	--	0.00	--	0.00	--
0.90	0.25	0.85		0.35	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.90	0.25	0.75		0.33	0.00	0.01	--	0.00	--	0.01	--	0.00	--
0.90	0.25	0.65		111.01	1.13	1.17	-0.03	1.24	-0.10	0.99	0.13	1.18	-0.04
0.90	0.25	0.55		117.46	1.20	1.03	0.14	1.17	0.02	0.92	0.23	1.17	0.02
0.90	0.25	0.45		1.75	0.02	0.01	--	0.01	--	0.01	--	0.01	--
0.90	0.25	0.35		0.49	0.01	0.00	--	0.00	--	0.00	--	0.00	--
Mean absolute error						0.04		0.03		0.05		0.02	
Average run time (minutes)						0.58		582		52		10	

Table 5-5: Simulated flux at the intermediate flux planes for day 58 of the multiple-source PCE dissolution experiment.

Day 58				Observed PCE Concentration mg/L	Calculated Flux mg/(cm <sup>2</sup> day)	SA		SCE		MRE		SA-MRE	
X	Y	Z				Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference	Simulated Flux mg/(cm <sup>2</sup> day)	Percent Difference
1.10	0.25	0.85		0.39	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.10	0.25	0.75		0.38	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.10	0.25	0.65		102.04	1.04	0.91	0.12	0.90	0.14	0.89	0.14	0.93	0.11
1.10	0.25	0.55		81.33	0.83	0.80	0.04	0.87	-0.05	0.81	0.02	0.84	-0.01
1.10	0.25	0.45		1.52	0.02	0.01	--	0.01	--	0.00	--	0.01	--
1.10	0.25	0.35		0.45	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.00	0.25	0.85		0.37	0.00	0.00	--	0.00	--	0.00	--	0.00	--
1.00	0.25	0.75		0.30	0.00	0.01	--	0.00	--	0.00	--	0.01	--
1.00	0.25	0.65		106.07	1.08	1.06	0.02	1.12	-0.04	0.94	0.13	1.05	0.03
1.00	0.25	0.55		112.76	1.15	0.92	0.20	1.09	0.05	0.91	0.21	1.00	0.13
1.00	0.25	0.45		1.86	0.02	0.01	--	0.01	--	0.00	--	0.01	--
1.00	0.25	0.35		0.51	0.01	0.00	--	0.00	--	0.00	--	0.00	--
0.90	0.25	0.85		0.35	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.90	0.25	0.75		0.33	0.00	0.00	--	0.00	--	0.00	--	0.00	--
0.90	0.25	0.65		111.01	1.13	1.11	0.02	1.14	-0.01	0.97	0.14	1.10	0.03
0.90	0.25	0.55		117.46	1.20	0.99	0.18	0.98	0.18	0.88	0.27	1.17	0.02
0.90	0.25	0.45		1.75	0.02	0.01	--	0.00	--	0.00	--	0.01	--
0.90	0.25	0.35		0.49	0.01	0.00	--	0.00	--	0.00	--	0.00	--
Mean absolute error						0.04		0.03		0.06		0.02	
Average run time (minutes)						0.62		587		55		10	



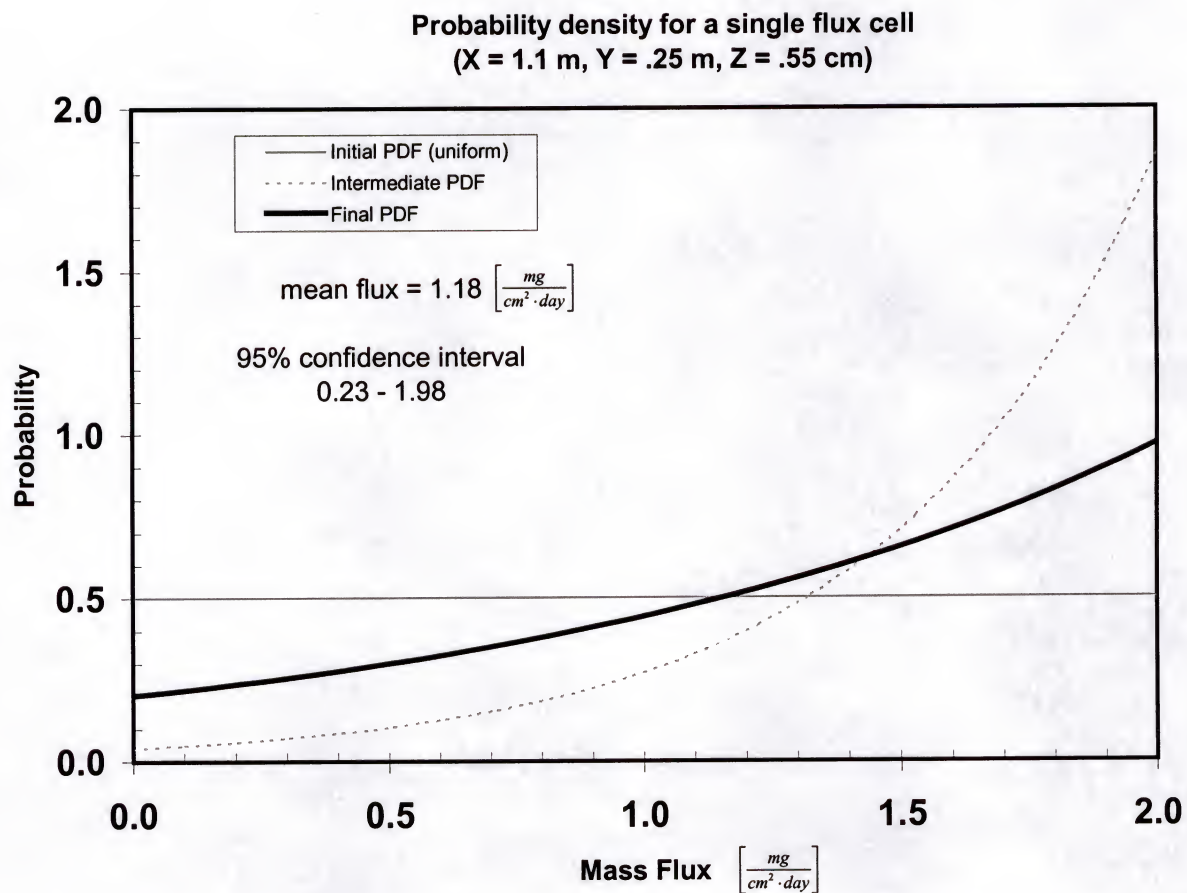


Figure 5-15. Estimated mean value, probability density function, and 95% confidence interval for a single flux cell from the PCE dissolution experiment.



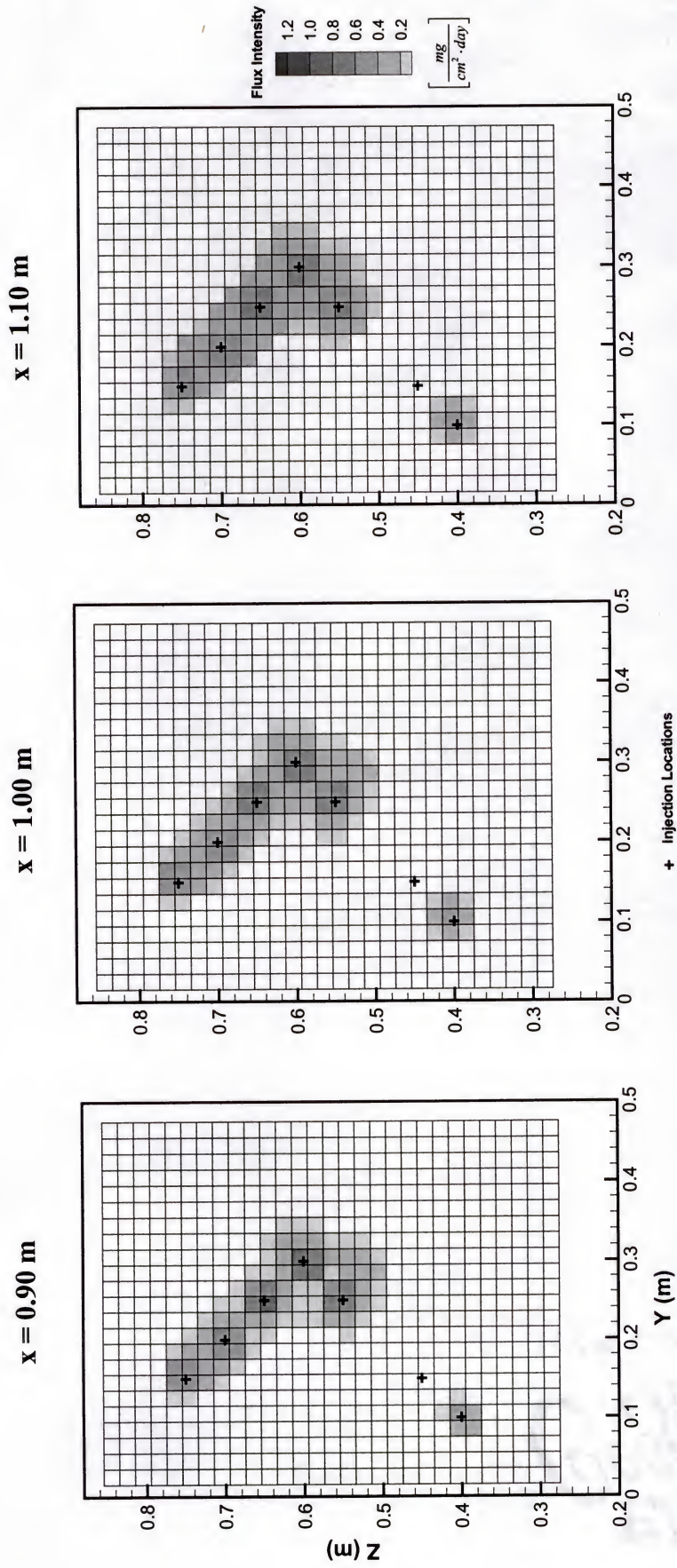


Figure 5-16. SA-MRE simulated PCE distributions at intermediate flux planes (Day 44).

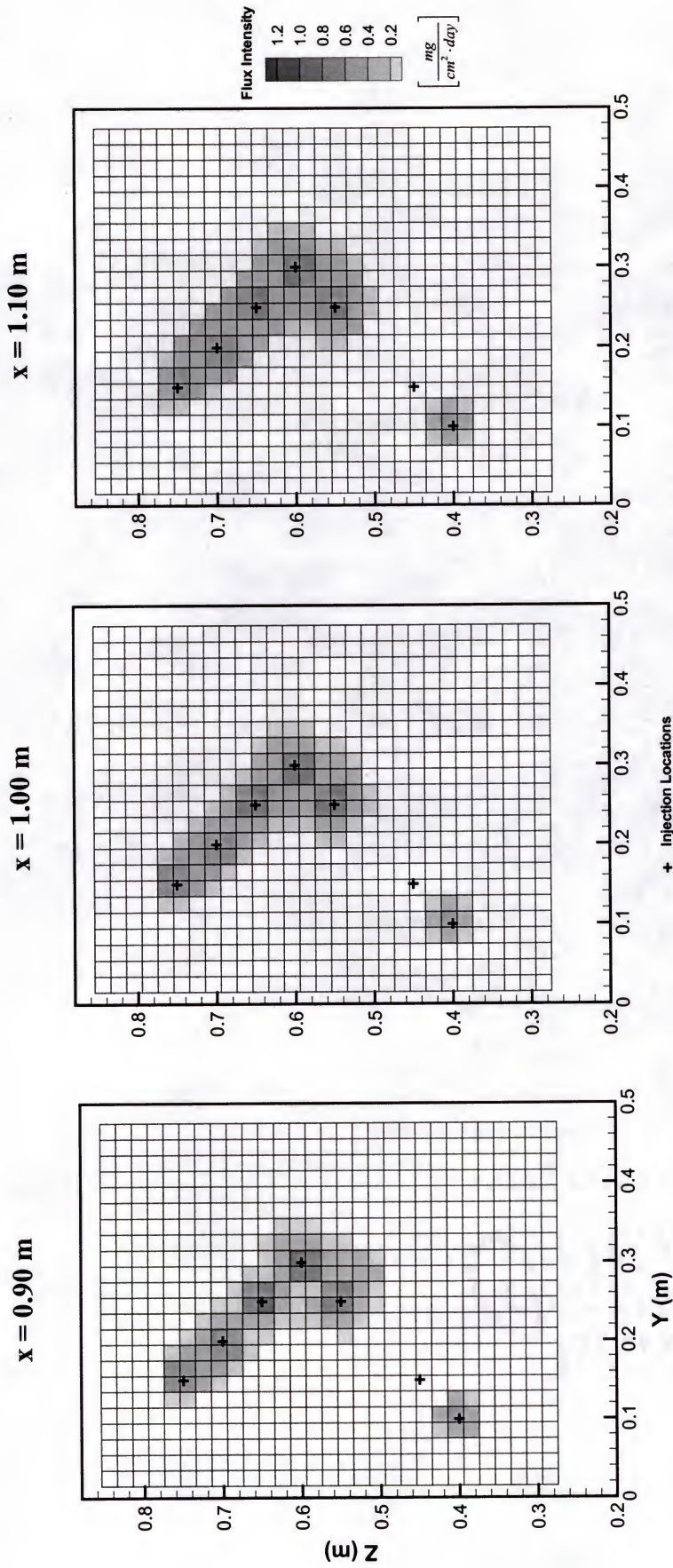


Figure 5-17. SA-MRE simulated PCE distributions at intermediate flux planes (Day 58).



the source zones increases (moving from  $x = 90$  m to  $x = 1.1$  m) the flux intensity at the central portion of the plume decreases while the lateral extent of the plume increases.

Figures 5-16 and 5-17 also demonstrate a critical relationship between the transverse dispersivity and the downgradient sampling strategy. It is known that NAPL solution was injected at location  $y = 0.15$  m,  $z = 0.45$  m and yet none of the models indicated any significant mass flux at this location. The cause for this error is two-fold: there were no concentration observations within the direct downgradient path of the source zone and the small order of magnitude of the system transverse dispersivity ( $\alpha_T = 0.0002$  m) did not allow for surrounding observations to detect any significant contribution from the source. So, even with the extensive data collected within the physical model, a known source was not detected due to the order of magnitude of the transverse dispersivity, which caused a gap in the downgradient sampling network.

Finally, each of the models was used to estimate the flux magnitude and distribution at the two source planes ( $x = 0.8$  m and  $x = 0.5$  m). The model simulated flux values are listed in Tables 5-6 and 5-7 for day 44 and day 58 respectively. The simulated flux values were verified by comparison to the effective PCE solubility range determined using Raoult's law (4-10) and a pure phase PCE solubility range of 150 – 200 mg/l. As discussed in Chapter 4, the resulting effective PCE solubility range is 105 – 140 mg/l, which corresponds to an expected experimental range for PCE flux of 1.07 – 1.43 mg / cm<sup>2</sup> day .

For source plane  $x = 0.8$  m all of the models predict source flux intensities that fall within the effective solubility range. For source plane  $x = 0.5$  m, the models predict flux intensities within the effective solubility range for the upper two source zones.



Table 5-6. Simulated flux at the source planes for day 44 of the multiple-source PCE dissolution experiment.

Effective PCE solubility		Day 44							
Upper Limit (mg/L)	Lower Limit (mg/L)	Observation coordinates				SA Simulated Flux mg/(cm <sup>2</sup> day)	SCE Simulated Flux mg/(cm <sup>2</sup> day)	MRE Simulated Flux mg/(cm <sup>2</sup> day)	SA-MRE Simulated Flux mg/(cm <sup>2</sup> day)
140	105	X	Y	Z					
		0.50	0.15	0.75		1.25	1.28	1.08	1.21
		0.50	0.20	0.70		1.28	1.38	1.27	1.36
		0.50	0.25	0.65		1.89	1.53	1.71	1.71
		0.50	0.30	0.60		1.97	1.60	1.72	1.67
		0.50	0.25	0.55		1.54	1.31	1.13	1.54
		0.80	0.25	0.65		1.29	1.39	1.28	1.40
		0.80	0.30	0.60		1.27	1.43	1.27	1.44
		0.80	0.25	0.55		1.01	1.29	1.23	1.30
		0.80	0.15	0.45		0.00	0.01	0.01	0.01
		0.80	0.10	0.40		0.53	0.67	0.63	0.74

Calculated effective flux		Average run time (minutes)				51	10
Upper Limit mg/(cm <sup>2</sup> day)	Lower Limit mg/(cm <sup>2</sup> day)					0.45	579
1.43	1.07						

Table 5-7. Simulated flux at the source planes for day 58 of the multiple-source PCE dissolution experiment.

Effective PCE solubility	
Upper Limit (mg/L)	Lower Limit (mg/L)
140	105

Calculated effective flux	
Upper Limit	Lower Limit
mg/(cm <sup>2</sup> day)	mg/(cm <sup>2</sup> day)
1.43	1.07

Day 58				SA	SCE	MRE	SA-MRE
Observation coordinates				Simulated Flux mg/(cm <sup>2</sup> day)	Simulated Flux mg/(cm <sup>2</sup> day)	Simulated Flux mg/(cm <sup>2</sup> day)	Simulated Flux mg/(cm <sup>2</sup> day)
X	Y	Z					
0.50	0.15	0.75		1.30	1.24	1.34	1.24
0.50	0.20	0.70		1.32	1.38	1.17	1.34
0.50	0.25	0.65		1.92	1.67	1.81	1.76
0.50	0.30	0.60		1.80	1.64	1.60	1.72
0.50	0.25	0.55		1.66	1.35	1.18	1.51
0.80	0.25	0.65		1.24	1.31	0.98	1.36
0.80	0.30	0.60		1.22	1.49	1.32	1.40
0.80	0.25	0.55		1.05	1.30	0.90	1.25
0.80	0.15	0.45		0.00	0.01	0.01	0.01
0.80	0.10	0.40		0.51	0.61	0.57	0.75

Average run time (minutes)	0.45	583	56	11
----------------------------	------	-----	----	----



But all of the models predict PCE flux intensities that exceed the effective solubility range by as much as 37% at the three source locations that are present on both the  $x = 0.5$  m and  $x = 0.8$  m source planes. These values are highlighted in Tables 5-6 and 5-7. It appears that because the sources are coincident on the two source planes, the flux plane model overcompensates for the contribution of the sources at the farther  $x = 0.5$  m source plane. This is most likely caused by the fact that because the sources are coincident on two consecutive source planes, the downgradient plume extents (based upon the transverse dispersivity) are greater than those expected for a single source at  $x = 0.5$  m. When attempting to estimate the flux at source plane  $x = 0.5$  m, the model has no method for taking into account the existence of another source downgradient at  $x = 0.8$  m. So, in order to match the downgradient plume extents, the model is forced to overcompensate the contribution of the source by increasing the flux intensity.

Figures 5-18 and 5-19 show the SA-MRE simulated flux distribution plots for day 44 and day 58 respectively. For comparison, the true source zone distributions recorded during excavation are displayed in Figure 5-20. The excavation data are represented as a scatter plot, where each point represents a point within the source zone as determined from a digital photograph. The colors (red or pink) correspond to the recorded color of dye within the source zone at that location. It is assumed that the red points indicate a PCE flux at or near the solubility limit, while the pink points may indicate a flux less the solubility limit.

As expected based upon the simulated flux values (Tables 5-6 and 5-7), the simulated flux distributions at  $x = 0.8$  m appear to (Figures 5-18B and 5-19B) compare well with the excavation source distribution (Figure 5-20B). The source located at



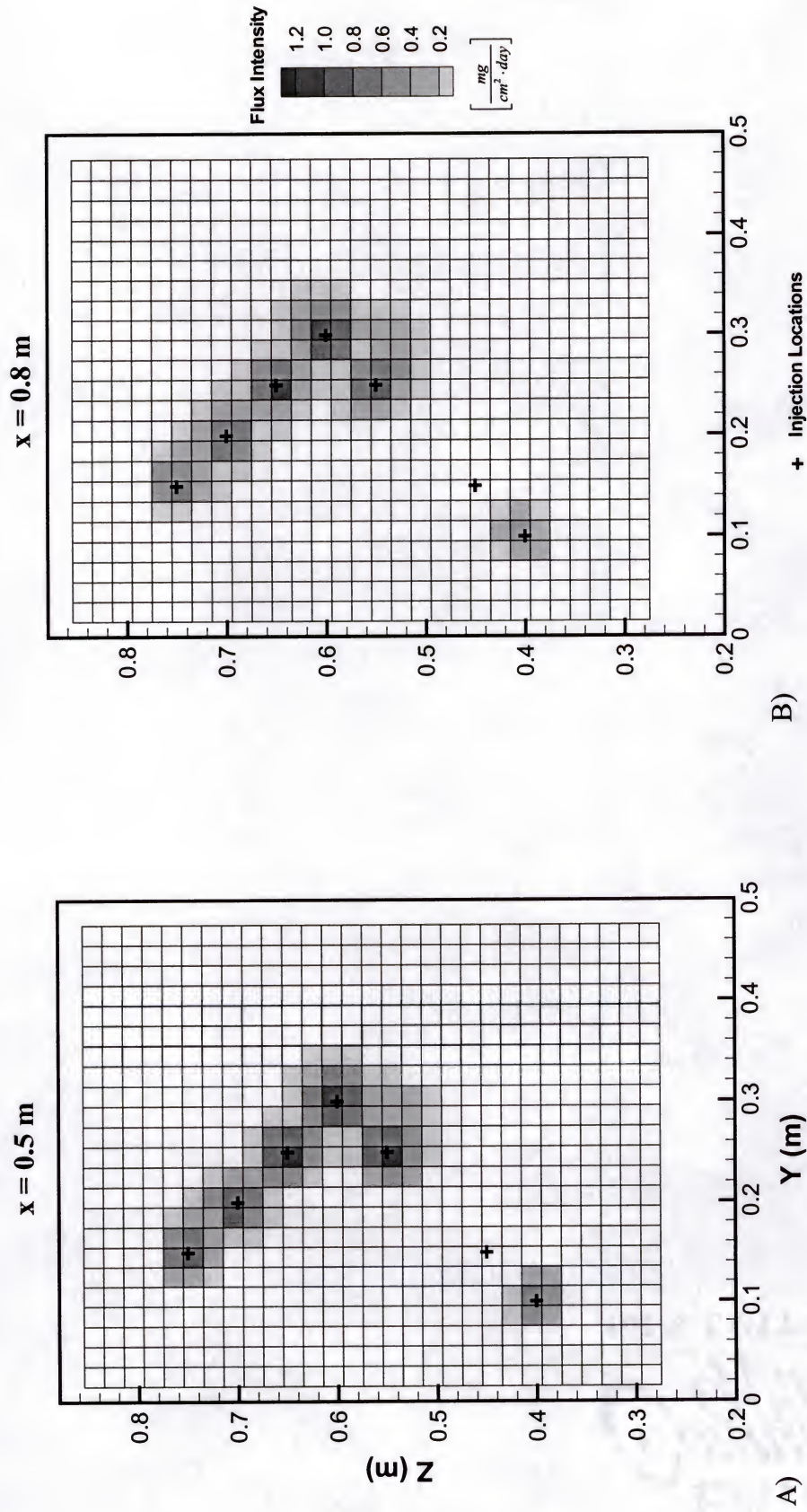


Figure 5-18. SA-MRE simulated flux at the PCE source planes (day 44).

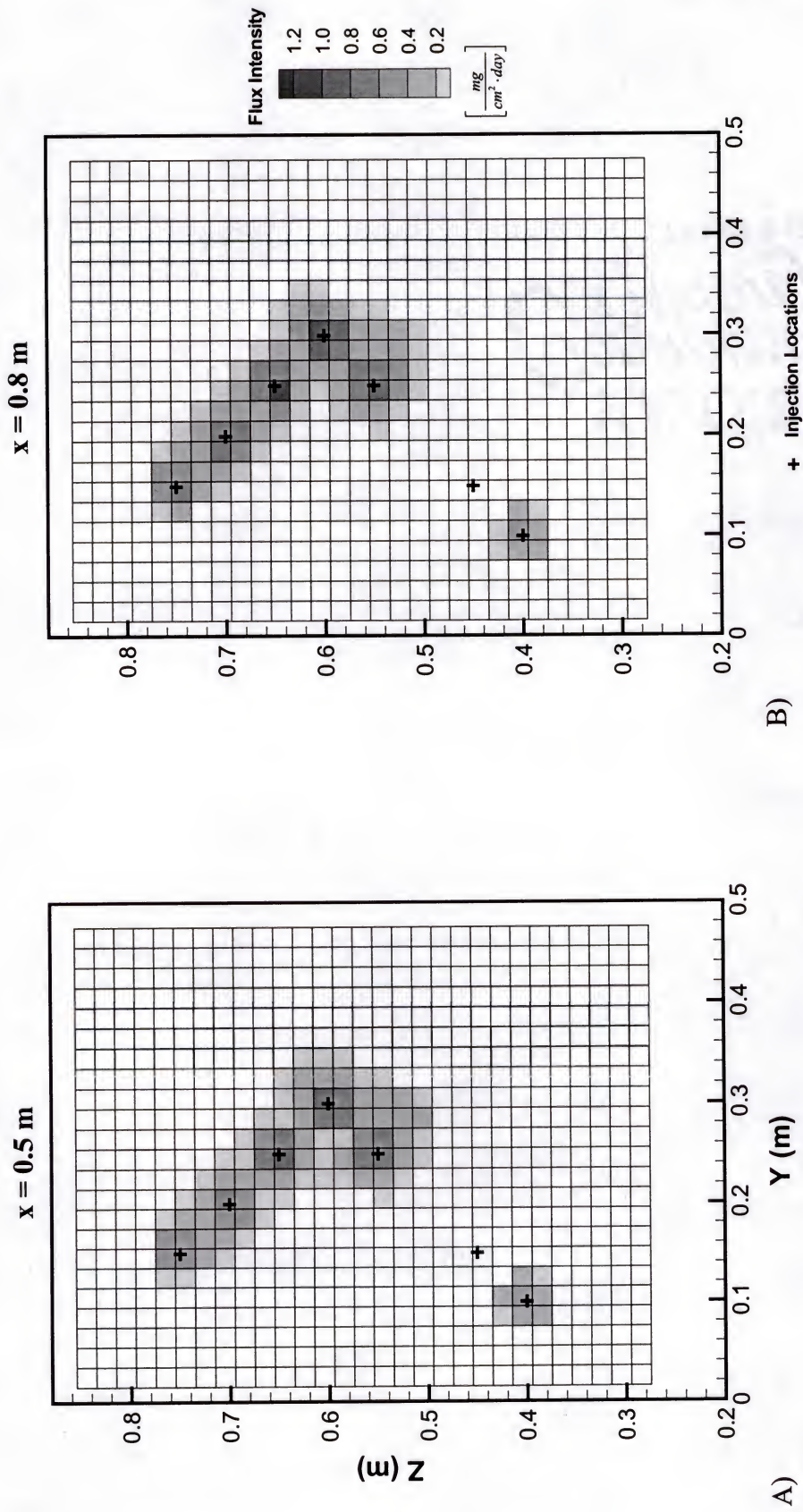


Figure 5-19. SA-MRE simulated flux at the PCE source planes (day 58).



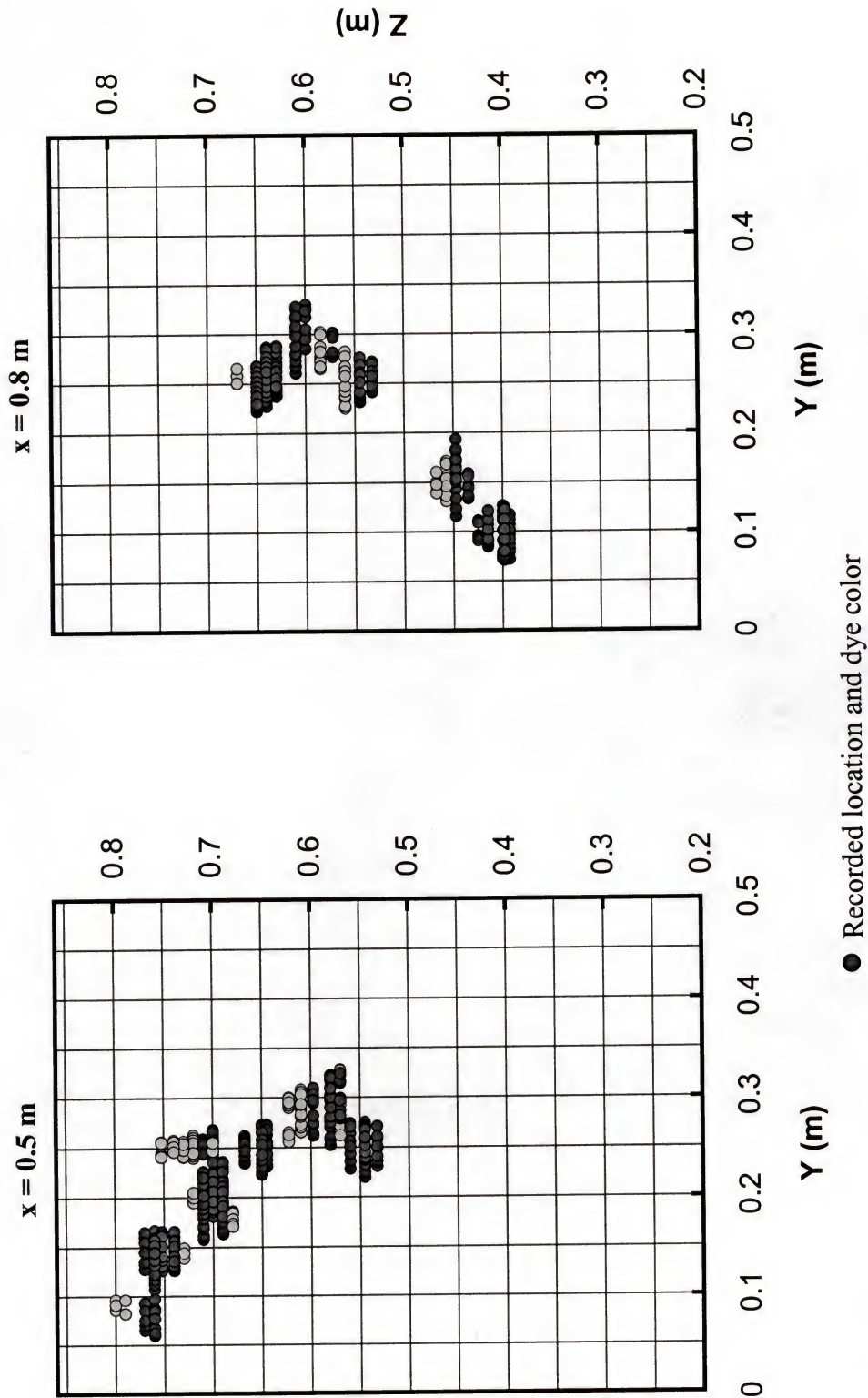


Figure 5-20. DNAPL source zone distribution recorded during excavation.



$y = 0.15$  m,  $z = 0.45$  m is not detected, as discussed previously, but elsewhere the simulated distribution seems to be relatively accurate. At flux plane  $x = 0.5$  m another complication due to the multiple source planes is evident. The flux plane model falsely indicated the existence of a source zone ( $y = 0.1$  m,  $z = 0.4$  m) which actually only existed at source plane  $x = 0.8$  m. This is not a surprising result, because the flux estimates for each source plane are determined independently there is no mechanism for the flux plane model to distinguish between sources that exist at the downgradient ( $x = 0.8$  m) source plane but not at the  $x = 0.5$  m source plane. The results could be improved by obtaining concentration data between the source planes, but this would require some knowledge of the source zone distribution to begin with. These results indicate that the model will perform well when estimating the flux magnitude and distribution at an intermediate flux plane or at a location near the down gradient edge of a contaminant source zone. However, it is evident that model simulated results are not as reliable if the flux plane is located within the extent of the source zone.

Application of the coupled SA-MRE algorithm allows for the reliability of the model simulated flux values to be estimated. Figure 5-21 shows the estimated 95% confidence limits for the farthest intermediate flux plane  $x = 1.1$  m and the source plane  $x = 0.8$  m. For both flux planes, all simulated mean flux values greater than  $0.01$   $\text{mg}/(\text{cm}^2\text{day})$  are shown sorted in descending order with their corresponding 95% confidence interval. The flux values not shown in the figure (below  $0.01$   $\text{mg}/(\text{cm}^2\text{day})$ ) represent the flux in the region outside of the contaminant plume. Because all observed PCE concentrations were below  $0.1$   $\text{mg}/\text{l}$  in this region, each of the models predicted simulated flux values below  $0.01$   $\text{mg}/(\text{cm}^2\text{day})$  with 95% confidence intervals of  $0.0$  to

0.03 mg / (cm<sup>2</sup>day) . These values were not shown in the figure in order to focus attention on the model-simulated values displaying greater uncertainty.

The simulated flux values greater than 1.0 mg / (cm<sup>2</sup>day) represent locations that were at or near the center of the contaminant plume, while values below 0.5 mg / (cm<sup>2</sup>day) represent locations that were at the lateral extents of the plume. Figure 5-21 demonstrates that as we move from the farthest intermediate flux plane ( $x = 1.1$  m) toward the source plane ( $x = 0.8$  m) the reliability increases (the 95% confidence intervals decrease) for the flux estimates located at the center of the plume and the lateral extents increases. However, it can be seen that for some of the simulated flux values within the range 0.5 to 1.0 mg / (cm<sup>2</sup>day) the reliability decreases at the source plane (the 95% confidence intervals increase). This is most likely due to a lack of concentration observations downgradient of these locations capable of verifying the model-simulated values. Figure 5-21 demonstrates that as we move from the farthest intermediate flux plane to the source plane the flux intensity at the central portions of the plume increases while at the lateral extents the flux intensity decreases. This is to be expected as we move closer to the source zone, and there is sufficient observation data to support the model-simulated values at these locations. However, at the intermediate locations, there may not be enough observation data to verify the model-simulated flux values resulting in greater uncertainty and larger 95% confidence intervals.

In summary, the application of the flux plane model with the three numerical solution techniques SA, SCE, and MRE has demonstrated these key points: SA is by far the most computationally efficient algorithm, a coupled SA-MRE algorithm utilizes the robust search capabilities of SA while providing the ability to estimate the uncertainty

associated with each of the simulated flux values, and that the flux plane model appears to be a capable tool for estimating contaminant mass flux at intermediate flux planes or at a location close to the downgradient edge of a contaminant source zone. The accuracy of the flux plane model is diminished if the flux plane is located within the extents of the contaminant source zone.



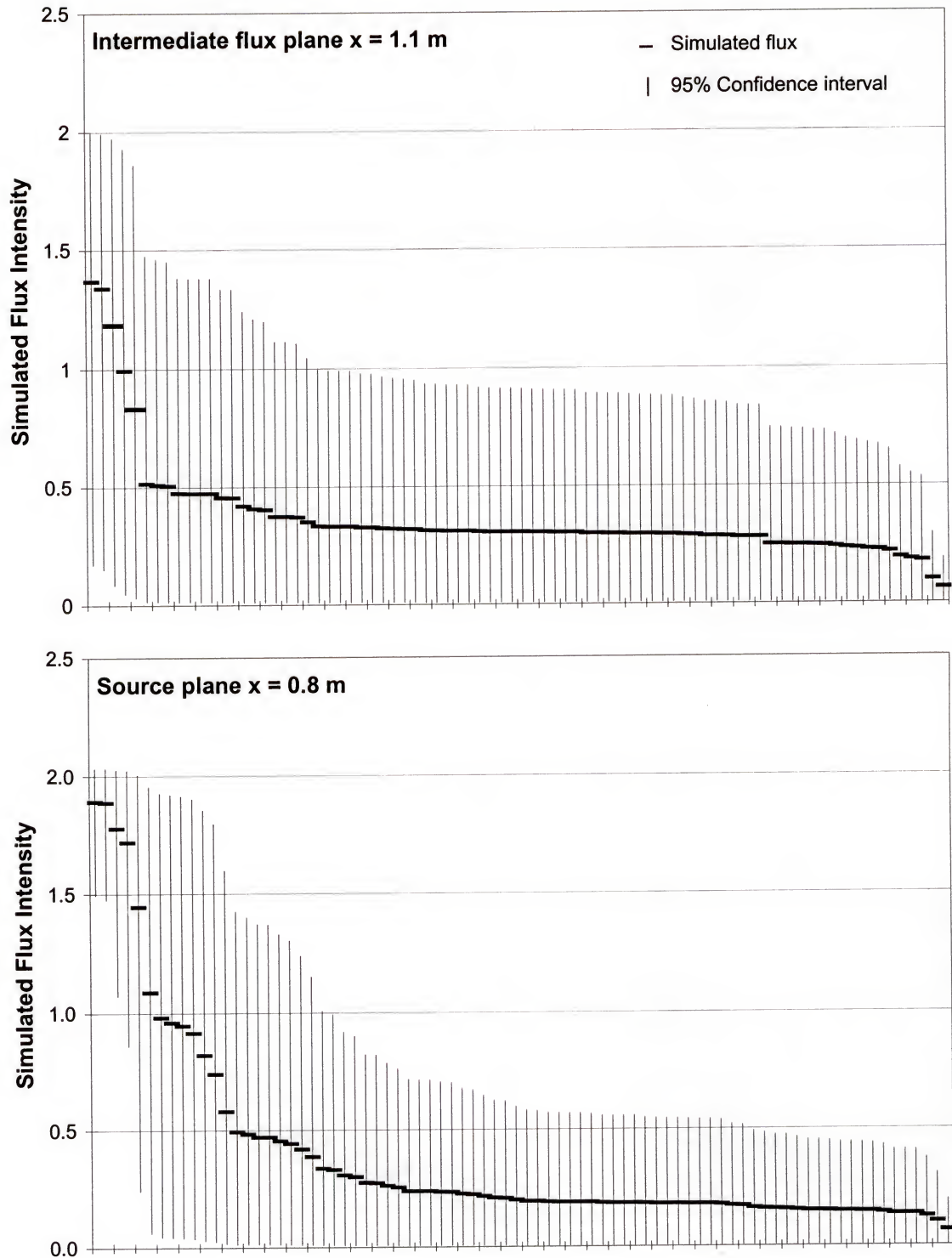


Figure 5-21. Simulated mean flux values and 95% confidence intervals. All flux values greater than 0.01 are shown sorted in order of descending flux intensity.

## CHAPTER 6

### SUMMARY AND CONCLUSIONS

The mathematical model presented in Chapter 2 was developed in order to estimate the flux magnitude and spatial distribution at an arbitrary flux plane based upon observed downgradient contaminant concentrations. The flux plane can be located at or near the source zone, or anywhere between the observation locations and the source zone. This means that the model can be used to estimate mass flux magnitude and spatial distribution at or near the source zone, or at some arbitrary boundary—such as the property boundary between a contaminated site and an adjoining property.

Each of the numerical solution techniques presented in Chapter 3 and applied in Chapter 5 are capable of solving multivariate nonlinear optimization problems. Simulated annealing (SA) and shuffled complex evolution (SCE) were considered because of their ability to solve nonlinear problems without requiring objective function derivatives. Minimum relative entropy (MRE) was considered because of its ability to provide estimates for parameter probability density functions and confidence intervals. While not as robust as the random search techniques, the ability to estimate parameter uncertainty makes minimum relative entropy a valuable method.

The flux plane model developed in Chapter 2 was solved independently using each of the numerical solution techniques and validated using experimental data from a three-dimensional aquifer model. Each solution technique was capable of estimating the magnitude and spatial distribution of mass flux through an arbitrary flux plane and the results were all essentially the same. The primary method for differentiating between the

independent algorithms was computational efficiency. Simulated annealing was by far the most robust and computationally efficient algorithm. Shuffled complex evolution provided results that were as good and at times better than simulated annealing, however the computational runtimes were excessive. Minimum relative entropy provided estimates that were comparable with SA and SCE with runtimes that were longer than SA but considerably less than SCE.

A coupled simulated annealing and minimum relative entropy solution technique was developed in order to take advantage of the robust solution capabilities of simulated annealing and the uncertainty estimation capabilities of minimum relative entropy. The coupled technique provided flux estimates that were similar to those of the independent methods. But, the probability density functions and confidence intervals provided by the coupled technique would not have been available from an independent SA algorithm and the coupled algorithm improves the computational efficiency of the independent MRE algorithm. The coupled algorithm provides a robust nonlinear solution technique with the ability to quantify the uncertainty associated with each unknown parameter.

The objective of this study was to develop a tool capable of assisting in characterizing contaminated groundwater sites and performing impact assessments. The flux plane model with the coupled SA-MRE entropy solution algorithm provides such a tool. This tool is by no means an all-encompassing groundwater transport model. In its current form, the model is only applicable under the conditions used to develop the mathematical model (uniform, steady-state groundwater flow conditions within a saturated unconfined aquifer assuming no retardation or degradation). As such, the model is a simple method for providing initial estimates for the magnitude and spatial



distribution of contaminant mass flux based upon observed groundwater contaminant concentrations. The model can be used to estimate mass flux intensity at or near a contaminant source zone, or at a property boundary. The model can be applied at any location of interest by simply establishing a flux plane perpendicular to the groundwater flow direction. The plane should be discretized so that the flux elements are equal to or smaller than the assumed source zone characteristic length. Because the model utilizes average hydrodynamic parameters and bulk hydrogeological property values, the model does not require a large amount of site characterization data. All that is required is an estimate of the groundwater velocity, a bulk estimate for porosity, an estimate of the magnitude of transverse dispersivity, and observed contaminant concentrations located downgradient of the flux plane. All of the requisite data are standard information obtained during a site contamination assessment. Using these data in conjunction with the flux plane SA-MRE model could be used to provide an initial estimate of the spatial distribution and magnitude of contaminant mass flux at a source zone, or crossing a specified boundary. This information could then be used in the planning process for additional site characterization, impact analysis, and future remediation.

## LIST OF REFERENCES

Aarts, E. and Korst, J. (1989). Simulated annealing and Boltzman machines: a stochastic approach to combinatorial optimization and neural computing. Chichester, England, John Wiley & Sons.

Agyei, E.O. (1997). Groundwater modeling and management using the finite element method and evolutionary optimization techniques. Dissertation, University of Adelaide.

Ahlfeld, D.P., Mulvey, J.M. and Pinder, G.F. (1986). Designing optimal strategies for contaminated groundwater remediation. *Advances in Water Resources* 9(June): 77-84.

Annable, M.D., Jawitz, J.W., Rao, P.S.C., Dai, D.P., Kim, H. and Wood, A.L. (1998a). Field evaluation of interfacial and partitioning tracers for characterization of effective NAPL-water contact areas. *Ground Water* 36(3): 495-502.

Annable, M.D., Rao, P.S.C., Hatfield, K., Graham, W.D., Wood, A.L. and Enfield, C.G. (1998b). Partitioning tracers for measuring residual NAPL: field-scale test results. *Journal of Environmental Engineering* 124(6): 498-503.

Bear, J. (1972). *Dynamics of fluids in porous media*. Mineola, Dover Publications.

Bear, J. (1979). *Hydraulics of groundwater*. New York, McGraw-Hill, Inc.

Bear, J. and Verruijt, A. (1987). *Modeling groundwater flow and pollution*. Dordrecht, Holland, Reidel Publishing.

Carrera, J., Medina, A., Heredia, J., Vives, L., Ward, J. and Walters, G. (1989). Parameter estimation in groundwater modeling: from theory to application. Groundwater contamination: use of models in decision-making, Amsterdam, The Netherlands, Kluwer Academic Publishers.

Domenico, P.A. and Robbins, G.A. (1985). A new method of contaminant plume analysis. *Ground Water* 23(4): 476-485.

Domenico, P.A. and Schwartz, F.W. (1990). *Physical and chemical hydrogeology*. New York, John Wiley and Sons.

Duan, Q., Sorooshian, S. and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* 28(4): 1015-1031.

Duan, Q., Sorooshian, S. and Gupta, V.K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology* 158: 256-284.

Duan, Q.Y., Gupta, V.K. and Sorooshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications* 76(3): 501-521.

Freyberg, D.L. (1988). An exercise in ground-water model calibration and prediction. *Ground Water* 26(3): 350-360.

Holland, J.H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, University of Michigan Press.

Hunt, B. (1978). Dispersive sources in uniform groundwater flow. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers* 104(HY1): 75-85.

Jin, M., Delshad, M., Dwarakanath, V., McKinney, D.C., Pope, G.A., Sepehrnoori, K., et al., (1995). Partitioning tracer test for detection, estimation, and remediation performance assessment of subsurface nonaqueous phase liquids. *Water Resources Research* 31(5): 1201-1211.

Kapur, J.N. and Kesavan, H.K. (1992). *Entropy optimization principles with application*. San Diego, Academic Press.



Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983). Optimization by simulated annealing. *Science* 220: 671-680.

Laarhoven, P.J.M.v. and Aarts, E.H.L. (1987). *Simulated annealing: theory and applications*. Dordrecht, Holland, D. Reidel Publishing Company.

Lesage, S. and Brown, S. (1994) Observation of the dissolution of NAPL mixtures. *Journal of Contaminant Hydrology*, 15: 57-71.

Liggett, J.A. (1994). *Fluid mechanics*. New York, McGraw-Hill.

Mabey, W.R., Smith, J.H., Podoll, R.T., Johnson, H.L., Mill, T., et al., (1982) Aquatic fate process data for organic priority pollutants. United States Environmental Protection Agency, Washington DC. Report No. 440/4-81-014.

Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.* 11: 431-441.

McLaughlin, D. and Townley, L.R. (1996). A reassessment of the groundwater inverse problem. *Water Resources Research* 32(5): 1131-1161.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21(6): 1087-1092.

Nelder, J.A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal* 7(4): 308-313.

Pankow, J.F. and Cherry, J.A. (1996). *Dense Chlorinated Solvents and other DNAPLS in groundwater*. Portland, Waterloo Press.

Petrucchi, R.H. (1985) *General chemistry principles and modern applications*. New York, Macmillan Publishing Company.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). Numerical recipes in FORTRAN: the art of scientific computing. Cambridge, Cambridge University Press.

Price, W.L. (1987). Global optimization algorithms for a CAD workstation. *Journal of Optimization Theory and Applications* 55(1): 133-146.

Rogers, L.L., Dowla, F.U. and Johnson, V.M. (1995). Optimal field-scale groundwater remediation using neural networks and the genetic algorithm. *Environmental Science and Technology* 29(5): 1145-1155.

Shore, J.E. and Johnson, R.W. (1981). Properties of cross-entropy minimization. *IEEE Transactions on Information Theory* IT-27(4): 472-482.

Sun, N.-Z. (1994). Inverse problems in groundwater modeling. Boston, Kluwer Academic.

Wagner, B.J. and Gorelick, S.M. (1987). Optimal groundwater quality management under parameter uncertainty. *Water Resources Research* 23(7): 1162-1174.

Wagner, B.J. and Gorelick, S.M. (1989). Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design. *Water Resources Research* 25(10): 2211-2225.

Wang, P.P. and Zheng, C. (1998). An efficient approach for successively perturbed groundwater models. *Advances in Water Resources* 21(6): 499-508.

Wolfson, R. and Pasachoff, J.M. (1990). Physics: extended with modern physics. Glenview, IL, Scott, Foresman and Company.

Woodbury, A.D. and Ulrych, T.J. (1993). Minimum relative entropy: forward probabilistic modeling. *Water Resources Research* 29(8): 2847-2860.

Woodbury, A.D. and Ulrych, T.J. (1996). Minimum relative entropy inversion: Theory and application to recovering the release history of a groundwater contaminant. *Water Resources Research* 32(9): 2671-2681.

Zheng, C. and Wang, P. (1996). Parameter structure identification using tabu search and simulated annealing. *Advances in Water Resources* 19(4): 215-224.



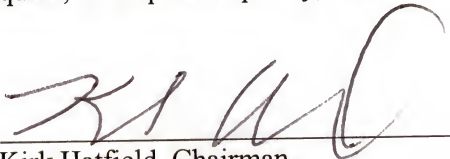
## BIOGRAPHICAL SKETCH

Mark Newman was born in Norwood, Massachusetts, on June 29, 1968. He graduated as Valedictorian of the graduating class of 1986 from Farmington High School, New Hampshire, and went on to attend Wentworth Institute of Technology, Boston, where he studied architecture for two years before transferring to the University of Florida, Gainesville. While studying civil engineering at the University of Florida he participated in the cooperative employment program, which allowed him the opportunity to gain work experience as both a field engineer and assistant estimator for the general contracting firm Danis Industries.

In 1993 he was married to Tricia Riskowitz, earned his Bachelor of Science in Civil Engineering, and began graduate school at the University of Florida. For the period of May 1995 through August 1996 he and Tricia lived in Panama City, Florida, where he performed a portion of his Ph.D. research at Armstrong Laboratories at Tyndall Air Force Base.

In August 1996, he and Tricia returned to the University of Florida so that he could complete his coursework and work towards completion of his dissertation. During the following years, he earned practical research experience while performing a groundwater field experiment at the Canadian Forces Base Borden, Ontario (Summer 1997), and a groundwater modeling study of Jacksonville, Florida, for the St. Johns River Water Management District (1988-1999).


I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



---

Kirk Hatfield, Chairman  
Associate Professor of Civil and Coastal  
Engineering


I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



---

P. Suresh C. Rao, Cochairman  
Lee A. Rieth Chair & Distinguished Professor  
Civil Engineering, Purdue University

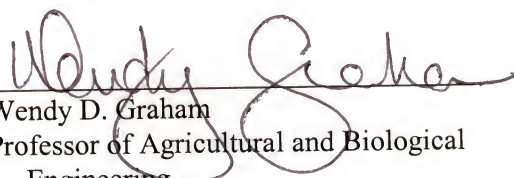
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



---

Michael D. Annable  
Associate Professor of Environmental  
Engineering Sciences


I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



---

Wendy D. Graham  
Professor of Agricultural and Biological  
Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

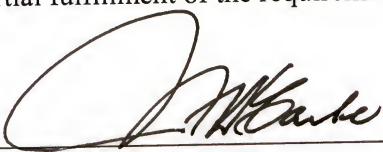


---

Jennifer M. Jacobs  
Assistant Professor of Civil and Coastal  
Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2001

  
Pramod P. Khargonekar  
Dean, College of Engineering

\_\_\_\_\_  
Winfred M. Phillips  
Dean, Graduate School